THE EVOLUTION OF GALACTIC NUCLEI

Mark Edward Bailey

B.A. (Cantab), M.Sc. (Sussex)

Presented for the Degree of Doctor of Philosophy

at the University of Edinburgh,

September, 1978

This thesis was composed by me and consists entirely of my own work

September, 1978

ABSTRACT

Observations of galaxies showing signatures of recently finished nuclear activity are reviewed, and it is shown that the available evidence supports the hypothesis that nuclear activity in massive galaxies is a recurrent phenomenon. This fact can be rather simply understood on a model in which stellar mass loss from the inner bulge components of galaxies ($r \lesssim 1$ kpc, say) flows inwards to form a dense large-scale nuclear disc, whose evolution then leads to periodic phases of nuclear activity. The structure and evolution of the disc are considered, and it is shown that the temperature is determined mainly by the heating effect of infalling matter. Even if this gas should be hot ($T \sim 2 \times 10^7$ K, say), the disc density is sufficiently high that its equilibrium temperature remains quite low. Close to the centre most of the gas is probably in molecular form.

The disc grows slowly in mass until the onset of a star-forming gravitational instability. At this point, strong turbulence driven by massive early-type stars leads to a rapid inward (viscous) transport of matter, which in turn leads to the formation of an active nucleus. In the first instance the disc forms a low-mass ($M \lesssim 10^6\ M_\odot$) low-entropy spinar, and the evolution of this object might in principle explain even the most luminous quasars. The late evolution of the system depends crucially on the late evolution of the spinar, and the two possibilities (1) that the spinar undergoes gravitational collapse, and (2) that the spinar undergoes a disruptive nuclear explosion can not yet be distinguished theoretically.

The theory is based on parameters appropriate to our own Galaxy, and observations of our Galactic Centre indicate that it is in a post-active state. The apparent absence of a supermassive black hole

($M \gtrsim 5 \times 10^6\ M_\odot$) argues against the black hole hypothesis, and it is tentatively concluded that the spinar's late evolution results either in a disruptive nuclear explosion, or in disruption from a state more tightly bound than E_{nuc} ($\sim 0{\cdot}7\ \%\ Mc^2$) by processes involving still undiscovered physical laws. The detailed evolution of low-mass spinars in normal galactic nuclei is thus an important theoretical problem for future investigation.

CONTENTS

I INTRODUCTION

In recent years it has become apparent that a great many extragalactic objects exhibit, to various degrees, some form of violent, energetic activity associated with the galactic nucleus. The luminosities of some of these objects, lying in the broad range $\sim 10^{36}$ W for an intrinsically faint Seyfert galaxy, up to $\sim 10^{41}$ W for the most luminous examples of quasars and BL Lac objects (e.g. Weedman, 1976a, Rieke et al, 1976), are so large that in many cases they completely outshine the light of the galaxy in which they are embedded. The associated energies are no less difficult to grasp - often exceeding by orders of magnitude the total rotational energy of an ordinary galaxy like our own - and range from $\sim 3 \times 10^{53}$ J for the typical Seyfert phenomenon (assuming a total lifetime of order 10^8 y and a typical luminsoity $\sim 10^{38}$ W, e.g. Weedman, 1976a), up to $\sim 10^{55}$ J for the more energetic examples of radio galaxies (e.g. Burbidge, 1971; de Young, 1976). These peak observed energies and luminosities (corresponding to $\sim 5 \times 10^7\ M_\odot c^2$ and $\sim 2 \times 10^{14}\ L_\odot$ respectively) are so large, that when quantities of these sizes were first indicated by observations, astronomers' initial reactions were to seek some unusual or exotic explanation for the data. A literature survey, covering the history of the first twenty-five years or so of the radio-galaxy/quasar/Seyfert-phenomenon, shows that a very wide range of hypotheses have been advanced, including, for example, colliding galaxies (Baade & Minkowski, 1954); collisions of matter with anti-matter (Burbidge, 1956); collisions of galaxies with extragalactic material (e.g. Shklovsky, 1962); delayed expansions, creation of matter and 'new physics' (e.g. Novikov, 1964; Ne'eman, 1965; McCrea, 1964; Ambartsumian, 1971); newly-formed, or young, galaxies (e.g. Field, 1964); and processes associated with superdense star clusters (e.g. Spitzer & Saslaw, 1966; Sanders, 1970; Arons et al, 1975). However the accumulation of more and better observational material has in recent years led slowly to a growing consensus amongst most astronomers that all forms of extreme activity shown by galaxies can be understood most simply as different manifestations of the same basic phenomenon which occurs, perhaps under different circumstances in different galaxies, in the nuclei of ordinary galaxies (e.g. Ozernoy, 1974; Rees, 1974; Rowan-Robinson, 1977; Weedman, 1976a,b). Quite apart from the merit of tying together (qualitatively) a large volume of diverse observational material, adoption of such a view, since it is implied that nuclear activity is a phenomenon connected in some way with the evolution of normal galaxies, has the advantage that theories of nuclear activity are forced to consider the whole range of new constraints imposed by observations of non-active galaxies. Within this framework, which is the one under-lying the arguments of the present thesis, theory must successfully answer two important interrelated questions: First, observed nuclear activity ought to be understandable in terms of a detailed physical model explaining luminosities, timescales, sizes and energetics; and secondly, it must be shown how this model fits into the overall scheme of normal galactic evolution. In particular (with reference to the Seyfert phenomenon), if nuclear activity occurs just once during a galaxy's evolution, theory ought to explain how activity can last as long as $\sim 10^8$ years, and - of equal importance - what switches it on (after a quiescent period of order 10^{10} y) and off again. On the other hand, if an individual Seyfert event should last for a much shorter time (10^5 - 10^7 y, say) then it is equally important to understand how different Seyfert events can occur sufficiently frequently that the cumulative duration of the phenomenon is still $\sim 10^8$ years.

The main emphasis of the present thesis is on the second question, namely how to explain the formation of an active nucleus as due to inevitable evolutionary processes occurring in normal galactic nuclei. However, because any theory which adequately explains the evolution up to an active state must place tight constraints on the sort of object likely to form at the centre of a galaxy, the arguments given here also have important implications for detailed models of the nature of nuclear activity. In the next chapters it is shown (1) that nuclear activity in typical massive galaxies like our own is probably a <u>recurrent</u> phenomenon, and (2) that this can be understood rather simply on a model in which stellar mass loss from the galaxy's bulge-component stars flows inwards to form a dense cold nuclear disc. Evolution of the disc leads naturally to the formation of a rapidly rotating low-entropy spinar of mass $M \lesssim 10^6 M_\odot$, and it is possible that most observed nuclear activity might be explained by the evolution of such an object. The apparent absence of a massive black hole in our Galaxy argues against the black hole theory.

The thesis divides broadly into three parts. The first, Chapters I - IV, consists of the present Introduction, a review of the observational evidence favouring recurrence (Chapter II), and a review of earlier theoretical attempts to understand the formation of active nuclei (Chapter III). Chapter IV summarises the arguments for favouring stellar mass loss as a source of nuclear fuel, and in particular shows that out-flowing supernova-driven galactic winds (Mathews & Baker, 1971) do not occur. The second part (Chapter V) consists of a detailed study of the structure of the disc prior to the formation of an active nucleus. This study shows that before the onset of gravitational instability, the dominant energy source in the disc is heating by infalling material. Even if this gas is hot ($\sim 2 \times 10^7$ K, say) cooling in the dense disc is sufficiently strong that its temperature can remain low, and it seems likely that most of the gas - especially that close to the centre - will be in molecular form. The third part (Chapters VI and VII) presents preliminary results of an initial investigation into the evolution of the disc, and arguments given there support the adoption of a spinar-like model for the source of energetic nuclear activity. Observations of the central region of our Galaxy can be easily understood in terms of the present theory, and in particular provide evidence that our Galaxy is in a post-active state. The thesis finishes with a short conclusion (Chapter VIII).

II OBSERVATIONAL EVIDENCE BEARING ON RECURRENCE IN SPIRAL AND ELLIPTICAL GALAXIES

It is possible, by comparing the relative numbers of active and non-active galaxies in a given class and assuming that nuclear activity *is* a phenomenon that occurs at some time or other in most normal galaxies, to estimate the cumulative duration of the active phase in that class of galaxy. For example de Vaucouleurs & de Vaucouleurs (1968) found that about 1 % of spiral galaxies showed Seyfert characteristics, implying a timescale for the Seyfert phase of order 10^8 years (assuming that spiral galaxies have ages $\sim 10^{10}$ y). In the case of the radio-galaxy phenomenon it has been shown that between $\sim$10 and 20 % of elliptical galaxies show detectable radio emission, and that this proportion and the mean radio luminosity both increase with the elliptical's absolute magnitude (e.g. Colla et al, 1975). The cumulative duration of the radio-galaxy phenomenon is thus of order 10^9 y, with the suggestion that some giant ellipticals may produce radio emission for a large proportion ($\sim$ 50 %, say) of their lives. It should be emphasised that these arguments, depending as they do on the relative numbers of active and non-active galaxies, can not be used to distinguish directly the hypotheses (A) that nuclear activity occurs just once during a galactic lifetime, and (B) that nuclear activity occurs as a number of distinct events. The latter phenomenon will be referred to as 'recurrence'.

The work of many investigators (e.g. Spitzer & Saslaw, 1966; Ozernoy & Usov, 1973; Hills, 1975) has usually been based simply on an *assumption* as to which hypothesis is correct. However the existence or otherwise of recurrence places strong restrictions on a great many detailed theories of the origin of nuclear activity. For example, a single extended period of activity lasting $\sim 10^8$ y would be hard to understand on the supermassive star concept (e.g. Fricke, 1973, 1974); equally, widely-spaced shorter-lasting periods of activity might be difficult to explain by a model involving the dynamical evolution of the nuclear star cluster (e.g. Sanders, 1970). For this reason a decision between the two possibilities (A) and (B) is of great importance, and in this chapter an attempt is made to compare their relative merits from an observational standpoint.

If nuclear activity <u>is</u> a transient phase through which most galaxies pass at one or more stages of their evolution, it is important to determine firstly what indications of recently finished activity might be shown by apparently normal galaxies, and secondly whether non-active galaxies showing the expected 'signatures' of previous activity have been observed. Provided that signatures of this kind can be identified reliably, an estimate of the number of active periods undergone by a particular class of galaxy can be obtained. This is done by comparing the estimated timescale for observability of a particular characteristic with the cumulative duration of the phenomenon as indicated by the fraction of galaxies showing the observational signature. Although it is not difficult to identify relevant observations, their correct interpretation requires some care, and for this reason the arguments of this chapter are set out in some detail. As will become clear, there are large uncertainties - both statistical and interpretive - and, as is often the case in astronomy, it is not possible to arrive at definite conclusions. However the present weight of evidence does seem to be in favour of recurrent nuclear activity, and the phenomenon of recurrence is therefore a possibility which merits serious theoretical investigation.

II-1 Evidence for recurrence in spiral galaxies

The Seyfert phenomenon, which is associated predominantly with the nuclei of spiral galaxies (e.g. Adams, 1977), is characterised by the following important features: (1) a compact highly luminous central object; (2) hot gas, with a large velocity dispersion; (3) outwardly moving gas clouds; and (4) strong radio emission from an extended central 'core' source. In addition, the Class 2 phenomenon is associated with the presence of large quantities of gas and dust (some of the gas being in molecular form, e.g. Rickard et al, 1977; Thompson et al, 1978); and a poorly-known fraction (between 10 and 60 %, Khachikian & Weedman, 1974; Adams, 1977) of Seyferts show indications of jets and filaments. It is interesting to note that the observation of jets associated with some non-active galaxies (e.g. NGC 1097; Wolstencroft & Zealey, 1975) might be indicative of recently finished activity; but the sample is small, and is not used here as evidence for recurrence.

The principal characteristic features of the spiral-Seyfert phenomenon listed above have led to various suggestions as to the most likely observable indications of recently finished nuclear activity (e.g. Lynden-Bell, 1969; Ozernoy, 1968; Van den Bergh, 1975; Van der Kruit, 1971; Lang & Terzian, 1969; Sanders & Bania, 1976). Following these authors it will be assumed here that past activity in a normal galaxy will manifest itself in one or more of the following ways:

(A) The existence of a massive non-stellar remnant.

(B) A burst of star formation in the nucleus.

(C) Non-circular velocities.

(D) Characteristic radio spectra and structure.

These possible indicators will now be discussed.

II-1A Massive non-stellar remnants

The interpretation of the evidence relating to this question depends to a great extent on the nature of the object that is thought to cause nuclear activity. For example, well-known arguments (e.g. Lynden-Bell, 1969; Rees, 1977) based on the possible collapse of massive evolved bodies into black holes have led to a growing body of opinion in which quasars and other active nuclei are phenomena caused by periods of rapid accretion onto massive black holes. This view implies that post-active galaxies ought to contain massive collapsed objects in their nuclei, and a number of suggestions have been made of ways to check this prediction (e.g. Wolfe & Burbidge, 1970). The hypothesis that a black hole represents the ultimate result (or cause) of a period of nuclear activity does however depend not only on the assumption that gravitational collapse occurs, but also that the collapse is correctly described by General Relativity. In general, since relativity theory has only been tested experimentally in the weak field limit, it is conceivable that a different sort of inert body might form (e.g. Opher, 1974), and it is even possible that an entirely different theory (e.g. Hoyle, 1971) might be the correct one to use. On the other hand, it has been shown by a number of investigators (e.g. Fricke, 1974, and references therein) that supermassive stars in a wide mass range may end their lives in disruptive nuclear explosions; and spinars too may have similar evolutionary end-points, or evolve otherwise so as not to leave a massive remnant (e.g. Flasar & Morrison, 1976). The nature and even existence of a remnant thus depends crucially on the theory of nuclear activity that is invoked, and consequently only model-dependent conclusions regarding recurrence can be obtained from this indicator. Unfortunately, because a remnant such as a black hole would normally be expected to survive for $\gg 10^{10}$ y with little essential change, even if a massive collapsed body was left behind after a period of nuclear activity, its detection in a normal galactic nucleus (for example by radiation from accreting gas) would not necessarily be directly relevant to the present question of distinguishing single activity from recurrence. The existence of such an object in a normal galaxy would imply simply that this galaxy had undergone at least one period of activity, with the precise number of active phases remaining undetermined. For this reason, although VLBI observations of many spiral galaxies do show evidence of the existence of mysterious objects in galactic nuclei (e.g. Kellermann et al, 1976, 1977), this possible indicator of recurrent nuclear activity is not conclusive.

Another kind of remnant, this time associated particularly with the evolution of a spinar-like object, might be the occurrence in a galactic nucleus of a dense cloud of cool gas (e.g. Ozernoy, 1968). However, although detection or otherwise of molecular clouds in normal galactic nuclei might be a good test of this kind of theory, it is arguable that such gas might have nothing at all to do with nuclear activity. Because of this, this indicator too is unsuitable for use in a general discussion of recurrence.

In conclusion, although the detection or otherwise of massive non-stellar remnants in the nuclei of normal galaxies is a useful means of checking the predictions of specific theories, the conclusions regarding recurrence from this indicator are not definitive.

II-1B Burst of star formation

This section divides broadly into three parts. First the theoretical argument for expecting a burst of star formation to accompany nuclear activity is summarised. This is followed by two estimates of the likely timescale for which such an event might remain observable; and lastly the available observational material is used to draw conclusions regarding the number of separate periods of nuclear activity.

A basic difficulty encountered in any discussion involving star formation is that the subject itself - although one of the most important in astrophysics - is still only poorly understood. However the problem of how to form stars out of an initially dispersed gaseous medium may be divided conveniently into two parts. First how to form, from the general interstellar medium, the clouds and cloud complexes which eventually produce stars; and secondly, how to understand fragmentation and the subsequent condensation of stellar-mass lumps within the major clouds. Here, although there are a number of theoretical difficulties -such as fragmentation, and how to rid the collapsing body of excess angular momentum - it is unlikely that star formation involves a significant fraction of the original cloud material; if this were not the case it would be difficult to explain the large amounts of gas remaining in our Galaxy at the present time. Returning to the problem of cloud formation, it has been noted (e.g. Tayler, 1976) that in the outer regions of galaxies the process of gravitational instability is likely in practice to be augmented by a number of potentially more rapid cloud-forming processes, such as thermal and hydromagnetic instabilities, and shocks caused by spiral arms, H II regions and supernovae. However in galactic nuclei in which star formation is not already occurring, not only are strong shocks likely to be virtually absent, but also the high star density in the nucleus inhibits gravitational instabilities until substantial masses of gas have accumulated. In the detailed theory presented in Chapters V and VI it is shown that the onset of gravitational instability and the subsequent generation of large mass motions driven by massive star formation in the nucleus leads naturally to central flows of matter high enough to explain a period of nuclear activity. It is concluded - on this theory at least - that star formation does accompany nuclear activity. If during its evolution the active nucleus ejects material and produces large non-circular motions (as are often observed in Seyfert galaxies) the associated shocks will very likely lead to further star formation in the surrounding disc, and consequently a post-active galaxy will be expected to show evidence of recent nuclear star formation. Observational support for the view that post-active galaxies are probable sites for bursts of star formation has been discussed for example by Van den Bergh (e.g. 1975, 1976).

In order to use the frequency of galaxies showing evidence of recent nuclear star formation as an argument for single or recurrent nuclear activity, it is necessary first to estimate the timescale for which such an event might remain observable. In the last part of this section reference will be made to observations of spiral galaxies with bright nuclear H II regions, and since these are emission nebulae excited by stars of spectral class earlier than B1 (Allen, 1973; p.259), the nuclear H II region phenomenon may well be observable for at least 10^7 y. If nuclear activity should trigger star formation simultaneously throughout the nuclear region of the galaxy, a lower limit of order 10^7 y for the duration of the phenomenon is indicated, but since it is probably more realistic to suppose that an initial burst of star formation near the nucleus would lead to secondary star formation further out as shocks due to expanding H II regions and supernovae ejecta propagate into the more distant neutral material, it is possible that the observable phenomenon might last somewhat longer. It is important to obtain an accurate upper limit to the duration of such an event, but without a detailed understanding of the full complexity of such a system (cf. Elmegreen & Lada, 1977), order of magnitude estimates will have to suffice.

One way to obtain such an estimate is to note that star formation is observed to occur in spiral arms. On the density-wave theory these are encountered at intervals on the order of 10^8 y in the outer parts of galaxies, and since the arms (as defined by H II regions and sites of star formation) are observed to be distinct from one another, it may be concluded that star formation in the outer regions of galaxies does not continue for times longer than a few x 10^7 y. If this is the case in the outer regions, then it _might_ also be true in the nucleus. This view is strengthened - at least in our Galaxy - by observations (see, for example, Moorwood, 1974) which indicate that conditions near the Galactic Centre are comparable to those in giant H II regions in spiral arms. A second estimate may be obtained from the requirement that an H II region, expanding at a few km s^{-1} and triggering star formation in any sufficiently dense material it happens to encounter, has time to expand beyond the immediate vicinity of the nucleus. By adopting an expansion velocity of order 5 km s^{-1} and assuming that the central region of the galaxy has a radius of $\sim$250 pc, a characteristic timescale of order 5×10^7 y is obtained. It is thus concluded that the observable duration of the nuclear H II region phenomenon lies in the approximate range $10^7 - 10^8$ years.

The proportion of spirals showing evidence of bright nuclear H II regions is very difficult to obtain accurately, since a complete sample of galaxies has not yet been surveyed with this statistic in mind. Markarian (1969) states that in a literature search of 600 galaxies brighter than the thirteenth magnitude more than 100 were found to have anomalies in their nuclear colour or spectral characteristics, but unfortunately he neither lists the galaxies in his sample nor describes in detail the nature of their anomalies. If the anomalies are interpreted as due to a recent burst of nuclear star formation, and if it is assumed that this followed a period of nuclear activity, then it may be concluded that ~ 15 % of galaxies brighter than thirteenth magnitude show evidence for recently active nuclei. Markarian states that in his sample the majority of anomalous galaxies were spirals, so it may be concluded that at least 15 % of spirals show evidence of this kind of Seyfert indicator.

An alternative procedure is to consider the galaxies in Markarian's lists which show narrow nuclear emission lines. This method too is subject to uncertainty (due mainly to the incompleteness of Markarian's lists and to the lack of detailed morphological data concerning Markarian galaxies), but since there is good evidence that galaxies showing narrow nuclear emission lines are indeed undergoing a burst of nuclear star formation (e.g. Pastoriza, 1975), this particular link in the argument is more secure.

Discussions of the statistical properties of Markarian galaxies have been given by Sargent (1972), Huchra & Sargent (1973) and Huchra (1977), who have shown that these galaxies comprise ~ 10 % of all galaxies fainter than $M_p = -22$. About 85 % of the sample show narrow nuclear emission lines (~ 11 % are Seyferts). Since Markarian galaxies are selected by the observational criterion of an ultraviolet continuum, this implies that $\sim 8\ \%$ of all galaxies show evidence for recent star formation and an ultraviolet continuum. The distribution of morphological types amongst Markarian galaxies is generally similar to that of field galaxies (Huchra, 1977), so at least $8\ \%$ of normal spirals show evidence for a recent burst of nuclear star formation.

These two estimates of the fraction of narrow-line spiral galaxies imply that $\sim 10\ \%$ of spirals, with a realistic uncertainty that is probably less than a factor of two, show signs of a recent burst of nuclear star formation. The cumulative duration of the nuclear H II region phenomenon is thus $\sim 10^9$ ($\times\!\!\div$ 2) years, which exceeds the duration of an individual burst estimated above by a factor of order 10. This is regarded as good evidence that spiral galaxies typically undergo many such periods of activity during their evolution. Because it has also been argued that star formation is a probable indicator of recently finished nuclear activity, it is concluded that Seyfert activity in spirals is a recurrent phenomenon.

II-1C Non-circular motions

It has long been known that Seyfert galaxies are often associated with the presence of massive outwardly moving clouds of gas (e.g. Walker, 1968). One of the most direct consequences of the mixing of ejected material with pre-existing gas moving in presumably near-circular orbits is that the resultant gas motions are likely to contain large non-circular velocities. Noting that this characteristic is indeed observed in the inner regions of a number of nearby spiral galaxies (including our own), a number of astronomers have argued that this _is_ evidence of previous nuclear activity and have attempted to infer from the observations the kind of activity that must have occurred (e.g. Van der Kruit, 1971; Sanders & Prendergast, 1974). Unfortunately, because non-circular velocities might in principle have a number of separate causes (such as tidal interaction with a neighbouring galaxy or companion, motion in the gravitational field of a hypothetical bar, or motion near the inner Lindblad resonance in the density-wave theory of spiral arms), a unique connection between non-circular motions and nuclear activity is difficult to prove. Thus, although it seems inescapable that observed nuclear activity would give rise to non-circular motions, the conclusions of this section are inevitably weakened by the various alternative explanations for this signature that have been suggested.

Accepting that non-circular motions are indeed caused predominantly by nuclear activity, in order to use observations of such velocity fields in normal galaxies to distinguish single from recurrent activity, it is necessary firstly to estimate the timescale for which observable motions might persist. Sanders & Prendergast (1974) found that oscillatory motions of the swept-up material might continue for a very long time after the initial expulsion of matter, and it was argued that non-circular motions might remain observable for $\gtrsim 10^8$ years. An accurate estimate of the decay timescale of such motions is very difficult to obtain, since it depends sensitively on the detailed nature of the interstellar medium. Here it will be assumed that 10^9 y represents a reasonable upper limit to the observable duration of the phenomenon (cf. Sanders & Prendergast, 1974).

The fraction of spirals showing evidence of non-circular velocities is also very difficult to estimate, as again a complete sample of galaxies has not been observed with this question in mind. However about half of the ordinary spirals (i.e. excluding barred spirals and Seyferts) discussed by Burbidge (1970) have non-circular motions in their nuclei, and - if he is correct in stating that this is a pseudo-random sample, not selected with nuclear activity specifically in mind - such a signature must be an extremely common characteristic of normal spirals, with a cumulative duration of order 5×10^9 y. If Seyfert activity is the dominant cause of such velocity fields it may be concluded that recurrence is again more strongly indicated than single activity.

In this context attention should be drawn to the recent discussion of non-circular motions in our own Galaxy by Clube (1978), in which it is shown that a wide range of observational data is consistent with a general expansion of the whole system. According to Clube, the very presence of spiral structure is an indication of previous activity, and since material arms have a characteristic timescale only of order 10^8 y, the conclusion that nuclear activity is recurrent is very strong. The argument rests, however, on a rather specific model of the active nucleus, and is therefore not suitable for use in the present general discussion.

To conclude this section, the evidence of non-circular motions is in favour of recurrent nuclear activity provided that they are indeed indicative of previous activity. However because there are several alternative explanations of a non-circular velocity field, the conclusion must remain tentative. Further improvement of the argument requires more detailed investigation, both theoretical and observational, aimed at distinguishing the various possibilities.

II-1D Characteristic radio structure

The radio properties of Seyfert galaxies have been discussed by a number of authors (e.g. Van der Kruit, 1973; Ekers, 1975; Van der Kruit & Allen, 1976; de Bruyn & Wilson, 1978), who have found that although their luminosities and radio structures do not usually compare with those of powerful radio galaxies, they do radiate 10 - 1000 times as much radio power as ordinary spiral galaxies. It is natural to assume that this difference in radio properties is a consequence of the observed nuclear activity, and because relativistic electrons radiating at radio wavelengths have synchrotron half-lives $\lesssim 10^8$ y (depending on their energies and the magnetic field), it might be possible to use radio observations of galaxies as a means of distinguishing single from recurrent activity (e.g. Lang & Terzian, 1969; Van der Laan & Perola, 1969). Unfortunately, firm predictions regarding the time evolution of the radio luminosity and spectrum of a post-active galaxy are very model-dependent, being sensitive to details such as the time-dependence of the injection spectrum, diffusion losses, and the assumed magnetic field. It is not, therefore, possible to come to definite conclusions concerning recurrence based only on the radio properties of galaxies. Van der Kruit (1973) however has noted that the available evidence does seem to indicate that nuclear activity in spirals is a phenomenon which is recurrent on a $10^8 - 10^9$ y timescale.

An alternative means of investigating the question of single or recurrent activity is to consider the evidence provided by anomalous radio structures found in some spiral galaxies (e.g. Oort, 1974b; Sanders & Bania, 1976). de Bruyn (1977) has recently investigated the frequency of occurrence of anomalous radio 'arms' amongst spirals, and from an estimate of their likely observable timescale and the assumption that they are related in some way to nuclear activity, has concluded, albeit with large uncertainty, that the evidence is in favour of recurrence.

II-2 Evidence for recurrence in elliptical galaxies

Firm evidence for recurrent nuclear activity in elliptical galaxies is more difficult to obtain than for the case of spirals, mainly because the apparent lack of a significant interstellar medium in these galaxies means that the indicators of recent nuclear star formation and non-circular velocities can not be used. Van den Bergh (1975, 1976) has discussed observational evidence bearing on the association of star formation with possibly recurrent nuclear outbursts in elliptical galaxies, and it has long been known (Van der Laan & Perola, 1969) that the long radio lifetimes of elliptical galaxies ($\sim 10^9$ y) compared to the much shorter radiative lifetime of the relativistic electrons means that nuclear activity in these galaxy types must either be recurrent or must occur continuously or quasi-continuously for a period of order 10^9 years. Evidence for repeated outbursts in elliptical galaxies is provided by observations of aligned structures on several widely different scales (e.g. Readhead et al, 1978), but whether the proper interpretation of such data should be in terms of recurrence or simply as quasi-continuous activity remains uncertain. Such examples can not easily be used as evidence in support of one particular view, and since it is conceivable that the observable effects of nuclear outbursts in elliptical galaxies might overlap, it is not even clear that the hypotheses of recurrence and quasi-continuous activity should be confronted at all. Observations relating to single or recurrent nuclear activity in ellipticals can only be definitive in the context of detailed models explaining the particular observed phenomenon, and can not be used generally as support for either hypothesis. It should be emphasised however that the available data are not inconsistent with recurrence.

II-3 Conclusions

This chapter has discussed the observational evidence bearing on the question of single or recurrent nuclear activity, and has shown that the most important evidence for recurrence in spiral galaxies is provided by (1) bursts of nuclear star formation, and (2) non-circular velocities. Interpretive difficulties do not allow definite conclusions to be drawn from the evidence of radio emission (though this was weakly in favour of recurrence), and unless a particular model of the nature of nuclear activity is assumed, arguments based on the existence or otherwise of massive non-stellar remnants are inconclusive. An important point to emphasise is that spiral galaxies showing one indicator very often show one or more of the others (e.g. our own Galaxy), indicating that their respective causes are in some way connected. Taken as a whole, the evidence with regard to spiral galaxies points more strongly towards hypothesis (B) - recurrence - than (A), and it is concluded that theories of the origin of the Seyfert phenomenon ought in principle to be able to account for the possibility of repeated nuclear outbursts. The position with regard to ellipticals is less clear, and it is conceivable that the observations might be understood in terms of a theory based on a single extended period of activity lasting typically 10^9 years. However it should be emphasised that the evidence is equally consistent with recurrence, and - especially if nuclear activity in ellipticals has fundamentally the same nature as that in spirals - theories of the origin of the radio-galaxy phenomenon ought also to be capable of explaining possibly recurrent nuclear activity.

III REVIEW OF POSSIBLE NUCLEUS-BUILDING MECHANISMS

This chapter divides broadly into two parts, depending on whether the theories under discussion are (1) those in which nuclear activity is regarded as essential to the very existence of galaxies ('nuclei build galaxies'), or (2) of the more conventional kind in which nuclei grow due to inevitable evolutionary processes which occur in galaxies. It is shown that no currently viable theory exists in the first catagory; for this reason - and also because such theories require 'new physics' for their success - it is argued that theoretical attention should concentrate initially on the second type of theory, namely that in which galaxies build nuclei. In this class of theories, those which depend on the dynamical evolution of the stellar component are placed in severe difficulty by the observational possibility that nuclear activity is a recurrent phenomenon in the evolution of most normal galaxies. It is concluded that any theory capable of easily explaining this observation must be based on the assumption that nuclei form by the evolution of the gaseous component of the galaxy.

III-1 Nuclei build galaxies

A great weakness of theories of this type is that the observations are not really fully explained (in the same sense for example as the theory of stellar evolution can be said to explain the observations of stars). Instead the theories are usually of a descriptive nature; the observed phenomena being used both as evidence for the break-down of conventional physical laws and as data with which to restrict the 'new physics' that follows. However if this objection is overlooked, the theories have great strength as a class in that they might in principle overcome at a stroke a large number of still unsolved astrophysical problems. As McCrea once pointed out, the problem of how to form a galaxy out of a tenuous intergalactic medium may only be a problem because nature does not do it that way !

The first suggestion that galactic nuclei are regions where the operation of unknown physical laws can be observed seems to have been that of Jeans (1928). He argued that the explanations of spiral structure then prevalent were themselves so implausible that it was almost inescapable to invoke 'new physics'. He went on to propose that the centres of spiral nebulae were of the nature of singular points out of which matter was poured into our universe.

Before thirty years had elapsed, another astronomer found himself forced by the nature of the observations into arguing that galactic nuclei were regions of unknown physics. Motivated primarily by the 'missing mass' problem in clusters of galaxies, Ambartsumian (e.g. 1958, 1961, 1965, 1971) suggested that these systems had positive energy and were in the process of flying apart. In his later work he has concentrated more on the relation of nuclear activity to the expanding systems, and concludes that the evidence for nuclei having

having the ability to eject large quantities of matter is now indisputable. The ejected matter is seen in a variety of forms, both relativistic and non-relativistic, and it was argued that masses ranging from the relatively small amounts seen in jets, to masses of galactic size could be involved.

Neither Jeans nor Ambartsumian attempted to give a deeper explanation of the phenomenon; the latter even arguing that to do so would be premature, since a thorough understanding of the observational data should always precede theoretical modelling (Ambartsumian, 1965). However, in the absence of a deeper explanation such theories become vulnerable to the charge that they are purely descriptive, lacking in any real predictive power and therefore not amenable to 'disproof'. Such criticism does not of course prove the theory groundless, but it does indicate that further development is necessary. Such developments were proposed independently by McCrea (1964), Novikov (1964) and Ne'eman (1965), who each brought cosmological concepts to bear on the problem.

McCrea suggested a modification of the Steady-State hypothesis, and argued that the available observations indicated that if matter was created anywhere then it was in the nuclei of galaxies. On this scheme continual creation led simply to the growth of galaxies, although the occasional system was able to eject a fragment of itself, which then became the 'embryo' for a new galaxy. In this way new galaxies could fill the gaps left by the old ones as the universe expanded.

Novikov and Ne'eman both considered modifications of the Big-Bang cosmology, and suggested that quasars should be identified with delayed cores or 'white holes'. On this scheme, quasars and active nuclei are the results of unknown initial conditions in the early universe which caused some parts to start their expansion after others.

A third proposal along these lines was made by Stothers (1966), who used concepts from both the Steady-State and Big-Bang cosmologies. He suggested that creation in the Steady-State theory occurred only where matter was lacking (the opposite of McCrea's suggestion), and proposed that quasars were the manifestation of such creation, occurring in massive bursts analagous to 'little bangs'. In the light of present evidence linking quasars to events in galactic nuclei (e.g. Weedman, 1976b), it would appear that this hypothesis is no longer tenable.

Other suggestions that ought to be mentioned at this point include the possibilities that growth of matter in Hoyle & Narlikar's C-field ('creation') cosmology might be an unstable process that could lead to bright point-like objects resembling quasars in appearance (see Hoyle & Narlikar, 1966a,b,c), and that a collapsing body in another universe might reappear in our own (e.g. Hjellming, 1971).

There is now a great deal of evidence indicating that conditions at large redshifts are very different from those nearby, showing that the simplest versions of the Steady-State hypothesis are no longer tenable. Theories involving 'white holes' have recently come under strong attack by Eardley (1974), who showed that they could only exist in the present universe if their masses were enormous - far outside the normal range of galactic masses that had been proposed for them. For these reasons (and also for the more basic reason noted at the beginning of this section) it is concluded that no persuasive or convincing theory currently exists in this catagory.

III-2 Galaxies build nuclei

Theories in this catagory are to be preferred, as a class, to those just discussed, because they rely on known physical laws and the long-term operation of processes that are inferred, or observed, to occur in galaxies. In addition, observations in principle place tight constraints on theory, which is very helpful in deciding which among many is most likely to work out successfully. The work of a number of observers (e.g. Morgan et al, 1971) has shown that Seyfert galaxies can usually be classified on existing morphological schemes, devised initially for so-called normal galaxies, and that the majority seem to be spiral or lenticular galaxies (e.g. Adams, 1977). Whenever their nuclear masses have been estimated (e.g. Richstone & Morton, 1975) they are found to be very similar to those that are observed in other galaxies, and Seyferts of both classes seem to be quite normal in dynamics and gas content outside their disturbed nuclear regions (e.g. Lewis, 1972). It is therefore difficult to maintain that Seyferts are 'monsters' amongst galaxies (to use Ozernoy's (1974) phrase), and it is far more reasonable to assume that a Seyfert galaxy is simply a 'normal' galaxy that happens to be undergoing a period of intense nuclear activity. This, as was noted in the Introduction, is the stance that has been adopted in the present thesis, and theories of nuclear activity are placed under tight constraints by the requirement that they be workable under quite ordinary conditions in normal galactic nuclei. In spite of these restrictions, however, it is still not possible to limit the various nucleus-building mechanisms to one kind, and the present discussion is accordingly divided into two further subdivisions, depending on whether or not formation of an active nucleus takes place solely due to the dynamical evolution of the galaxy's stellar component.

III-2A Theories relying on stellar dynamical processes

A large number of theories of the central source of activity in galactic nuclei postulate that the activity is caused by processes such as stellar collisions which might occur in extremely dense star systems. However, because observations of normal galactic nuclei do not in general indicate sufficiently high densities, the presence of the required conditions must be either assumed, or shown to be a necessary consequence of the evolution of galaxies as they are now observed. In the former case the problem reduces essentially to the demonstration that, from a theory of the origin of galaxies, it is reasonable that a few galaxies should be formed initially with very dense nuclei; in the latter case a mechanism must be found by which the timescale for growth of a dense stellar core can be reduced to a value $\lesssim 10^{10}$ y. The observational evidence that active galaxies differ insignificantly from normal galaxies, except in respect of their activity, is a major difficulty for any theory advocating a 'primordial' origin for the required nuclear conditions, and if recurrence is accepted, the problem of finding a fast-acting process by which to change stellar orbits is made even more difficult.

The timescale on which the evolution of stellar orbits takes place may be written (e.g. Spitzer & Schwarzschild, 1951) in the form:

$$T_{ref} \sim \sqrt{\tfrac{2}{3}}\, \frac{1}{3\pi}\, v_s^3 \Big/ G^2 m^2 n \ln(N/2)$$

where v_s is the r.m.s. velocity dispersion, N is the total number of stars, m their mass, and n their space density. Spitzer & Saslaw (1966) showed that this could be re-written (for polytropes with indices in the range 0 to 4) in terms of the r.m.s. radius

R of the system as

$$T_{ref} \sim \frac{8{\cdot}3 \times 10^5 \, N^{1/2} \, R^{3/2}}{m^{1/2} \left(\log_{10} N - 0{\cdot}3\right)} \quad y \qquad (1)$$

where R is measured in pc and m is in solar units.

Dynamical evolution of this kind leads inexorably to the growth of a dense stellar core, and numerical calculations reported by Spitzer & Thuan (1972) indicated that in a one-component system the inner 10 % of the mass would collapse towards the centre in about 20 reference times. The evolution of systems containing stars of various masses is more complicated, owing to a tendency towards equipartition of energy. However it has been shown (e.g. Spitzer, 1969; Saslaw & de Young, 1971) that in realistic situations equipartition is not possible, and the heavy stars tend to fall towards the centre where they form a dense self-gravitating system with a large velocity dispersion. The timescale for the formation of a dense core by this process can be shorter than the corresponding timescale in a one-component system, and for a two-component cluster containing stars of masses m_1 and m_2 $(m_1 < m_2)$ it was shown to be smaller than the reference time by a factor of order m_1/m_2 (Spitzer & Hart, 1971). Under normal conditions it is likely that $m_1/m_2 \nless 0{\cdot}1$, so it may be concluded that dense stellar cores can probably not develop on timescales $\lesssim 0{\cdot}1$ reference times. The reference time for a 'typical' galactic nucleus, containing $\sim 10^{10}$ solar-type stars within a radius ~ 1 kpc (cf. our own Galaxy), can be estimated from the formulae above to be $\sim 3 \times 10^{14}$ y, so evolution by distant star-star 'collisions' is unlikely to be an important effect leading to the formation of a massive central object. The reference times for star systems with length scales much less than a kiloparsec are of course much reduced, but nuclear star clusters with densities in the ranges indicated by observation still either have very long relaxation times, or do not contain significant amounts of mass (see for example infrared observations of our own Galactic Centre).

An alternative way by which a nucleus might be formed has been proposed by Tremaine et al (1975), who have argued that the effects of dynamical friction on globular clusters (due to their motions through a galaxy of much lighter stars) would cause them to gradually spiral into the centre on a timescale on the order of 10^{10} years. However although this is a plausible way by which to increase the number of stars in the central regions of a galaxy, it does not seem that the process could lead to a significant reduction in the time-scale for growth of a dense stellar core.

Two main conclusions are drawn from this work. First, dynamical evolution by distant two-body gravitational encounters <u>does</u> lead to the growth of a dense stellar core. However for typical observed nuclei, the growth timescale is always much longer than the currently estimated ages ($\sim 10^{10}$ y) of galaxies, so in the absence of a detailed theory of the origin of galaxies explaining the exceptional 'initial conditions' required to lower these timescales significantly, the process does not seem to be capable of easily explaining the underlying normality shown by active galaxies. The second general conclusion is that all theories of nuclear activity which are based on a process involving only the dynamical evolution of the stellar component of a galaxy lead most naturally to <u>one</u> period of nuclear activity. The constraint of recurrence therefore virtually eliminates such hypotheses.

III-2B Theories relying on gas dynamical processes

One of the first to emphasise the important role that gas might play in explaining nuclear activity was Shklovsky (1960). He argued that an extremely high supernova rate might be expected during the earliest phases of galactic evolution, when star formation had only recently been initiated, and therefore identified radio galaxies with the early stages of the evolution of giant galaxies. Ginzburg (1960), in a similar study, pointed out that even the gravitational energy that was released when a massive cloud fragmented into proto-stars was sufficient to explain the energies of radio sources. Another who explicitly related quasars to the early stages of galactic evolution was Field (1964), who included not only supernovae, but also the luminosity due to the large numbers of massive main-sequence stars that were presumed to have been formed; and a number of other authors have related quasars to processes which might occur specifically at a very early stage of galactic evolution (e.g. Hoyle, 1961; Layzer, 1965; Piddington, 1964, 1966; Sturrock, 1966). An alternative model relating the formation of a nucleus to the earliest stages of a galaxy's evolution was proposed by Lynden-Bell (e.g. 1971), who argued that viscous evolution of a massive gaseous disc would lead inevitably to the formation of a massive black hole. With the exception of this more recent suggestion, all the above theories have severe difficulty in explaining continuing activity (i.e. nuclear activity associated with demonstrably old galaxies), and are clearly unable to explain recurrence.

Influenced by the difficulty of understanding continuing activity (in particular the association of many radio sources with elliptical galaxies composed of old stars), some astronomers considered alternative ways by which the gaseous component of a galaxy could give rise to observed activity. Cameron (1962), for example, discussed the problem of star formation in an elliptical galaxy, and concluded (if turbulence could be neglected) that gas in such a galaxy could not fragment into stars until its mass exceeded $\sim 3 \times 10^8 M_\odot$. In this way it was suggested that bursts of star formation could be associated with quite old galaxies, and the high supernova rate consequent upon such an event was put forward as an explanation of nuclear activity. Shklovsky (1962) extended his 1960 argument, and proposed that activity might be expected in otherwise very old galaxies if extragalactic material occasionally fell into the centres of elliptical galaxies.

An obvious source of material from which to build an active nucleus is the gas lost from evolving stars in the spheroidal component of the galaxy. Spitzer (e.g. 1971) considered this source in the context of the dense star cluster model of nuclear activity, and argued that the gas lost from stars in a spherical system would most likely collapse towards the centre and fragment into new stars, thereby forming a dense subsystem. This process could be repeated, and eventually a sufficiently dense stellar core would be created in which dynamical relaxation could take over, and lead in less than 10^{10} years to the formation of a core dominated by stellar collisions. Shklovsky (1971) also considered stellar mass loss as a possible source of material, but thought that instead of forming stars, the gas was most likely to form a supermassive magnetoid, or spinar.

About this time an important work was published by Mathews & Baker (1971), whose calculations showed, subject to a number of assumptions, that the gas lost from stars in a typical elliptical galaxy was most likely to be blown out of the system as a supernova-driven galactic wind. It was thus difficult to see how stellar mass loss could ever lead to nuclear activity. However if the supernova rate was too low, or the central gas density too high, cooling processes could dominate heating, and the wind was found to develop a thermally unstable core which eventually collapsed into the galactic centre. Mathews (1972) showed that this collapsing gas was unlikely to form stars, and argued instead that it would probably form a single coherent object which might then be the cause of a period of nuclear activity.

To conclude this chapter, the main difficulty with the hypothesis that the growth of an active nucleus is due to the dynamical evolution of the stellar component, is that stars move on orbits of nearly constant energy and the timescales required for change from observed states to ones sufficiently dense to give rise to a period of activity are prohibitively long. In contrast, gas motions are generally very dissipative, and for this reason gas might be expected to collapse into the nucleus on a relatively short timescale. The gaseous component as a source of fuel for nuclear activity thus has the great advantage that continuing and even recurrent nuclear activity might be readily explained. For these reasons it is concluded that the active nuclei of galaxies are probably caused primarily by evolution of the galaxy's gas.

IV REASONS FOR PREFERRING STELLAR MASS LOSS AS A SOURCE OF FUEL FOR NUCLEAR ACTIVITY

In the previous chapter it was noted that the hypothesis of recurrence places tight constraints on theories which explain the generation of a period of nuclear activity by normal processes of galactic evolution, and it was concluded that formation of an active nucleus is probably due to the evolution of the gaseous component of the galaxy. The gas out of which the nucleus is formed could conceivably originate in sources outside the main body of the galaxy (e.g. Shklovsky, 1962), but since very little is currently known even regarding such basic questions as the numbers, sizes and masses of intergalactic gas clouds, it is important first to consider the possibility that the galaxy itself is the primary source of its nuclear material. In this case there are two possible sources of gas: (1) gas left over from the epoch of galaxy formation, and (2) gas lost by stars undergoing normal stellar evolution. The first of these, because it is not directly observable, is the more uncertain, and estimates of the mass, density distribution and angular momentum of this gas as a function of time must rely on a detailed theoretical model of galaxy formation. Some of these models (e.g. Gott & Thuan, 1976) predict that gas is unlikely to survive for very long after the epoch of initial collapse in elliptical galaxies and the central regions of spirals, but because of unavoidably large uncertainties in input physics and astronomy the reliability of the calculations is difficult to assess. However in view of this result, and also because of the possibility that an early 'quasar' phase of galactic evolution might sweep primordial gas completely out of the nuclear regions (Wolfe, 1974), it is assumed here that by the time the galaxy formation process has been completed the central regions of a galaxy may be treated as essentially devoid of gas. In this way any theory that is developed is made relatively free of unknown initial conditions, and the discussion is restricted - at least initially - to a seemingly inevitable source of gas: material lost by stars undergoing normal evolution.

The problem of showing that stellar mass loss can lead to observed nuclear activity divides broadly into three main parts. First, it must be shown that the mass loss rate is sufficient not only to give rise to activity on a timescale shorter than 10^{10} years, but also to account for the energy requirements of a typical active nucleus. Secondly, it must be shown that the gas can lose energy and collapse towards the galactic centre; and lastly, it must be shown that the final collapse can occur at a sufficiently fast rate that a period of nuclear activity can be explained, either by formation of a massive object such as a spinar or by a period of rapid accretion onto a black hole. In practice this last condition requires inward flows at a rate in excess of $\sim 0{\cdot}1\ M_\odot\ y^{-1}$, and the difficulty of creating such large central flows of matter is one of the major problems that any theory of nuclear activity must overcome. This chapter deals with the first two aspects of the problem; the treatment of the structure and evolution of the disc formed by infalling gas is reserved for Chapters V and VI.

In order for galactic material to give rise to nuclear activity, it is necessary that it loses both energy and angular momentum. However, although the strong dissipation usually associated with the dynamics of interstellar gas clouds is itself a good reason for considering a gaseous source of material, it is not easy - even for gas - to lose angular momentum efficiently on a galactic scale. Indeed

Lynden-Bell & Pringle (1974) have shown that angular momentum transport in a gaseous disc with viscosity roughly equal to that expected for a cloudy interstellar medium is only important on timescales less than 10^{10} y for radii $R \lesssim 6$ kpc. Since a major aim of the present work is to investigate a theory capable of explaining <u>recurrent</u> nuclear activity, the timescales of interest are $\ll 10^{10}$ y, and the conclusion - that viscous evolution can affect significantly only the inner regions of a galaxy - applies in even stronger form. A recurrently active nucleus is thus only likely to form from low angular momentum gas residing in the inner few kiloparsecs of the galaxy. If the source of nuclear material is stellar mass loss, attention must therefore concentrate on the behaviour of the gas lost from the relatively low angular momentum stars comprising the bulge component of a galaxy.

IV-1 Calculation of the stellar mass loss rate

This section presents a calculation of the rate of stellar mass loss from a nuclear star cluster, and shows that the specific mass loss rate (mass loss rate per unit mass of stars) is given to within a factor of order 2 by

$$\alpha_* \sim 10^{-2}\, t^{-1} \quad y^{-1} \qquad (t \gtrsim 10^8\, y)$$

where t is measured in years. The cumulative fractional mass loss in 10^{10} y is $\sim 16\%$, of which just over half occurs at epochs later than $t \sim 10^8$ y. The section concludes with a short discussion of the uncertainties underlying these estimates.

It is assumed that star formation occurs simultaneously throughout the bulge-component of a galaxy at $t = 0$, and that the initial mass function (IMF) has the Salpeter (1955) form, defined by

$$\phi(m) = k\, m^{-(1+x)} \qquad m_L \leq m \leq m_U \tag{2}$$

where $\phi(m)$ is defined to be the fraction by number of stars with masses in the interval $(m, m+dm)$. Normalisation implies

$$\int_{m_L}^{m_U} \phi(m)\, dm = 1 .$$

$$\Rightarrow \quad k(m_L, m_U, x) = x \big/ \left(m_L^{-x} - m_U^{-x}\right) \tag{3}$$

The mass lost by one star during its evolution is defined to be $L(m)$, and it is assumed that the bulk of this occurs instantaneously at the

at the end of a star's evolutionary lifetime of length $\tau(m)$.
L(m) is taken to be the same as that used by Tinsley (1976):

$$L(m) = \begin{cases} 0.84\,m - 0.44 & 1 \leq m \leq 6 \\ m - 1.4 & m > 6 \end{cases} \qquad (4)$$

Here both L and m are measured in solar units; and Larson's (1974a) analytic approximation to $\tau(m)$ (based on the compilation of evolutionary lifetimes constructed by Tinsley (1972)) is also used:

$$\log(\tau(m)) = 10.02 - 3.57\log m + 0.90(\log m)^2 \qquad (5)$$

In this formula τ and m are measured in years and solar masses respectively. (5) implies

$$-\frac{d\tau}{dm} = \frac{\tau}{m}\left[3.57 - 1.8\log m\right] \qquad (6)$$

If the total number of stars formed initially is N, the number with masses between m and m+dm is N $\phi(m)$dm. Each of these injects a mass L(m) into the interstellar medium at $t = \tau(m)$, so the mass loss rate at time t is given by

$$\frac{d\mathfrak{M}}{dt} = \frac{N\,\phi(m)\,dm\,.\,L(m)}{\tau(m) - \tau(m+dm)}$$

$$\text{i.e.} \quad \frac{d\mathfrak{M}}{dt} = N\,\phi(m)\,L(m)\Big/\left(-\frac{d\tau}{dm}\right) \qquad (7)$$

where m is given as a function of $t = \tau$ by (5).

The initial stellar mass is

$$\mathcal{M}_o = N \int_{m_L}^{m_U} m\,\phi(m)\,dm$$

$$\Rightarrow \quad \mathcal{M}_o = \frac{Nk}{(x-1)}\left(m_L^{1-x} - m_U^{1-x} \right)$$

so the specific mass loss rate α_* , defined by

$$\alpha_*(t) = \frac{1}{\mathcal{M}_o}\frac{d\mathcal{M}}{dt}$$

is given by

$$\alpha_*(t) = \frac{(x-1)}{\left(m_L^{1-x} - m_U^{1-x}\right)} \cdot \frac{L(m)\,m^{-x}}{\left[3{\cdot}57 - 1{\cdot}8\log m\right]} \cdot \frac{1}{t} \qquad (8)$$

i.e.

$$\alpha_*(t) = \frac{k(m_L, m_U, x-1)\,.\,G(x,t)}{t} \qquad (9)$$

where k is a function (defined in (3)) depending only on the assumed initial mass function, and $G(x,t)$, defined below, has only a weak time-dependence.

$$G(x,t) = \frac{1}{\left[3{\cdot}57 - 1{\cdot}8\log m\right]} \cdot \begin{cases} 0{\cdot}84\,m^{1-x} - 0{\cdot}44\,m^{-x} & 1 \le m \le 6 \\ m^{1-x} - 1{\cdot}44\,m^{-x} & m > 6 \end{cases} \qquad (10)$$

(m is related to $t = \tau$ by (5)).

Further progress requires an assumption to be made as to the initial mass function. For definiteness it will be assumed here that the IMF has $x = 1{\cdot}35$ (e.g. Salpeter, 1955) and that the upper limit to its range is given by $m_U = 50$. The choice of the lower limit, m_L, is less easy, as not only is this difficult to determine observationally, but also compelling evidence for extending a mass function of the assumed form towards the lower mass range ($\lesssim 1\ M_\odot$) does not exist. This problem may be avoided, following Talbot & Arnett (1973), by the introduction of the parameter ζ defined to be the fraction of mass in the IMF which consists of stars more massive than the sun. It may be shown that m_L, m_U, x and ζ are connected by the relation:

$$\zeta = \left(1 - m_U^{1-x}\right) \Big/ \left(m_L^{1-x} - m_U^{1-x}\right)$$

Talbot & Arnett (1973) argue that a reasonable estimate of this quantity is $\zeta \sim 0{\cdot}25$ (so that $3/4$ of the mass is composed of stars with $m \leqslant 1$). Adopting this value here implies

$$m_L(m_U, x) \sim \left(4 - 3\, m_U^{1-x}\right)^{1/1-x} \qquad (11)$$

which, assuming $m_U = 50$ and $x = 1{\cdot}35$, gives $m_L \sim 0{\cdot}035$ and $k(m_L, m_U, x-1) \sim 0{\cdot}12$.

The slowly-varying function $G(x,t)$ is tabulated below, for $x = 1{\cdot}35$ and $5 \times 10^6 \leqslant t \leqslant 10^{10}$, and by combining these values with the estimate of k just made, the specific mass loss rate may be shown to be

$$\alpha_*(t) \sim 10^{-2}\, t^{-1} \qquad y^{-1} \qquad (12)$$

where t is in years. This agrees with the independent estimate of this quantity made by Gisler (1976).

Tabulation of the function defined in equation (9) :

t (y)	$m(\tau = t)$	G(1.35, t)
5×10^6	30·9	0·32
10^7	16·7	0·25
5×10^7	6·6	0·19
10^8	4·8	0·18
5×10^8	2·6	0·17
10^9	2·0	0·16
5×10^9	1·2	0·13
10^{10}	1·0	0·11

The cumulative fractional mass loss is given by

$$f(t) = \int_{t_{min}}^{t} \alpha_*(t)\, dt$$

Using the results obtained above and taking $t_{min} \sim 5 \times 10^6$ y it may be shown that $f(t = 10^{10}) \sim 16\ \%$, of which a little more than half occurs during times $t \gtrsim 10^8$ y. This result is in good agreement with that of Talbot & Arnett (1973), and it is concluded that stellar mass loss could provide on the order of 10 % of the initial mass of the cluster in $\sim 10^{10}$ years. In particular, since the mass of the bulge component of a galaxy like our own is $M_n \sim 10^{10}\ M_\odot$ (Sanders & Lowinger, 1972), stellar mass loss from this part of a massive spiral galaxy might provide up to $\sim 10^9\ M_\odot$ as possible fuel for nuclear activity.

The principal uncertainty underlying these calculations lies in the assumptions made as to the initial mass function, where, because the problem of star formation is not yet understood, the physical causes of a particular IMF remain obscure. Consequently it is not known whether the IMF ought to be regarded as a 'universal' function, essentially independent of time and position within the galaxy, or as a function whose properties in the distant past might have been very different from what they are now. There are some observational indications that the IMF might have been different in the past (e.g. Limber, 1960), but other authors (see Talbot & Arnett, 1973) have been able to account equally well for the observations in terms of a model with an unchanging IMF. The position regarding the universality or otherwise of the IMF is thus open, and it is possible that the mass loss rate calculated above might be subject to a systematic error. At the present time there is no way to avoid this difficulty, and the approach adopted here - to assume that the IMF of the bulge-component stars closely resembles that of stars now in the solar neighbourhood - was thought to be the best means of progressing. Observations of the local IMF (e.g. Salpeter, 1955; Talbot & Arnett, 1973) support a power-law approximation in the high-mass range ($M \gtrsim 1\ M_\odot$), with x in the range $1{\cdot}3 - 1{\cdot}8$, but in the low-mass range the evidence for such a mass distribution is not compelling (e.g. Larson, 1973; Tinsley, 1978). Fortunately, the exact form of the low-mass IMF is not of great importance in the present problem, since low-mass stars do not contribute to the mass loss rate significantly on timescales $\lesssim 10^{10}$ y. Because of this, the present approach, which characterised the low-mass IMF by the fraction $(1 - \zeta)$ of the mass consisting of stars with $m \lesssim 1$, is sufficiently accurate for our purpose. Given

$\zeta \sim 0{\cdot}25$, it may be verified that varying x in the broad range $1{\cdot}1 \lesssim x \lesssim 1{\cdot}8$ leaves the order of magnitude of $\alpha_*(t)$ unaltered. From this result it is concluded that the approximation (12) is probably accurate to within a factor of ~ 2.

IV-2 Discussion of the galactic wind problem

An important objection to the hypothesis that stellar mass loss fuels nuclear activity is the possibility that gas lost by stars in elliptical galaxies is driven out of the system as a hot supernova-heated galactic wind (Mathews & Baker, 1971). In view of the many similarities between elliptical galaxies and the bulge components of spirals, it is likely that a similar conclusion applies to stellar mass loss in the central regions of these galaxies too (Faber & Gallagher, 1976), and in this case - however indicative the arguments given in the previous chapter - stellar mass loss could not be invoked as the source of material for nuclear activity. It is therefore important to decide whether or not a galactic wind is likely to occur, and the main aim of this section is to show, by presenting a critical review of galactic wind theory, that there are several reasons for believing that galactic winds do not efficiently remove stellar mass loss from a galaxy.

The properties of steady-state galactic winds have been discussed by a number of authors, and in particular it has been shown (Johnson & Axford, 1971) that three basic types of steady flow are possible, depending on whether or not the system contains a sink. If there is no sink, steady flows must be entirely outwards; but if a sink is allowed, the flow can either be wholly inwards or partly inwards and partly outwards, depending on the sink's strength. Not all combinations of basic wind parameters however produce suitable conditions for a steady-state flow (for example the gas might be too cold to be able to escape from the galaxy), and a solution to the problem of whether or not steady flows are in fact likely to occur in galaxies requires a physically realistic treatment of the time-dependent problem. One such solution was given by Mathews & Baker (1971), who considered the evolution of stellar mass loss in a typical elliptical galaxy (NGC 3379) taking account of radiative cooling and heating by hot stars and supernovae. In order to make the problem tractable a number of physical assumptions and approximations were made, the most important of which were

(1) At $t = 0$ the galaxy was assumed to be devoid of gas.

(2) Material and energy were assumed to be deposited into the interstellar medium smoothly, both in space and time.

(3) No sink was allowed.

(4) Cooling of the hot gas by dust was neglected.

Not surprisingly, in view of assumption (3) above, whenever a steady wind was favoured it was found to be an outward flow; but the most important conclusion to be drawn from their work is that for 'reasonable' supernova and gas injection rates, normal elliptical galaxies could be expected to lose a non-primordial interstellar medium as an outflowing galactic wind.

It is important to emphasise that this conclusion does not apply during the early stages of galactic evolution; for example Larson (1974b) has shown that the much higher stellar mass loss rates during the first $\sim 10^9$ years of galactic evolution (cf. Section IV-1) led to a thermally unsteady flow. During this stage of galactic evolution stellar mass loss is not removed from the system, and instead the flow will resemble the thermally unsteady winds computed by Mathews & Baker (1971). In these cases, although the gas did initially heat up, its density became large enough for cooling to dominate heating before a wind could be produced. Once this had happened, the central regions (where the density was highest) cooled rapidly and collapsed into the origin, followed closely by the hotter more rarefied gas from further out which rushed in to take its place. This gas too became subject to thermal instability, and it was found that eventually a substantial fraction of the gas lost from stars was able to fall into the centre during a series of non-linear cycles.

The galactic wind calculations did not include the effects of angular momentum. However it is reasonable to suppose that if this had been taken account, then it would bave been found that the collapsing gas would have formed a dense rapidly rotating disc. Thus, during epochs $t \lesssim 10^9$ y (before thermally steady flows are expected to occur), stellar mass loss will flow inwards and the system will form a dense nuclear disc. What happens at later times depends crucially on the effects of the (possibly) hot infalling material on the disc, and in the next chapter it is shown that any disc that is likely to be formed will be dense, cold and highly dissipative. The disc is sufficiently dense to be able to absorb even hot infalling material without heating up, and it may therefore be concluded that the appropriate solution to the galactic wind equations at epochs $t \gtrsim 10^9$ y will be one including the presence of a massive disc-like sink. Provided that no other agency (such as nuclear activity or sweeping by an intergalactic medium) can destroy the disc and completely eliminate the gas from the central region of the galaxy, stellar mass loss - even at relatively late epochs ($t \sim 10^{10}$ y, say) - will continue to flow inwards.

Should this argument not be accepted, removal of stellar mass loss by an outflowing galactic wind is still not a necessary conclusion, and several alternative arguments can be constructed which question this assumption. First, although there is general agreement that Type II supernovae have massive progenitors, the evidence relating to the progenitors of Type I supernovae (which are the ones supposed to drive a galactic wind) is still ambiguous (Tammann, 1974, 1977). For example Tinsley (1977) has argued that the observations are quite consistent with the view that <u>all</u> supernovae have massive progenitors. If this was the case a galactic wind could only be a transient feature of galactic evolution, since - by removing the supply of gas from which to form new massive stars - it would be self-defeating. Secondly, it was assumed that dust cooling of the hot gas was negligible, but it has since been shown (Burke & Silk, 1974) that at typical wind temperatures ($T \sim 10^7$ K) dust cooling may in fact be dominant. Although the lifetimes of grains in such a hot environment are probably not long enough to significantly cool the gas, if evolving stars produce large quantities of dust the equilibrium dust fraction might still be an important coolant. Thirdly, in the case of a spiral galaxy, the bulk of the galaxy's mass lies outside the nuclear region. In this case, although conditions might be favourable for production of a 'nuclear wind' from the galaxy's bulge component, it is not certain that the gas would be removed permanently from the galaxy; it is possible that a fraction of the wind material might fall back onto the disc in the form of cool clouds, some of which might return to the nuclear region itself. An additional difficulty with extending the galactic wind hypothesis to spirals, as noted by Faber & Gallagher (1976), is that the existence of the disc greatly complicates the nature of the problem, and the flow might be altered in ways that can not be easily foreseen. Lastly, it seems unlikely in practice that supernovae ejecta would heat all the interstellar medium uniformly. If the gas lost from stars contains density inhomogeneities (as it almost certainly does; e.g. Capriotti, 1973), and if allowance is made for the non-uniform way that mass and energy are actually injected into the interstellar medium (assumption (2)), the physical state of the interstellar medium will be much more complex than that assumed in the numerical calculations. In this situation it could be argued plausibly that supernovae would give rise to a wind involving only a small fraction of the interstellar medium (perhaps $\lesssim 50\%$ by volume; Cox & Smith, 1974), leaving the remainder to cool and fall back towards the disc.

Thus, although at first sight the galactic wind hypothesis seems to have strong observational and theoretical support (Faber & Gallagher, 1976), there are several possible ways of avoiding the general conclusion. Of these the most important is linked to the assumption implicit in Mathews & Baker's work, that the system contains no sink; and it has been shown that current understanding of galactic evolution leads to the prediction that when the flow first becomes steady, the central regions of a galaxy will contain a massive disc-like sink. The disc can survive even hot infalling material (Chapter V), so the flow at later epochs will be one dominated by the presence of the disc. Although this is regarded as the most compelling argument against an outflowing wind, the above additional arguments show that the case <u>for</u> galactic winds still remains to be proven. It is concluded that stellar mass loss, certainly from stars in the central regions of galaxies ($r \lesssim 1$ kpc, say), probably does flow inwards, where it will form a dense gaseous disc in the nuclear regions. The structure and evolution of the disc are discussed in the next two chapters; but before turning to these questions it is first necessary to verify that stellar mass loss is a <u>sufficient</u> fuel for nuclear activity. This may be done by comparing the total energy requirement of the Seyfert phenomenon ($\sim 2 \times 10^6\ M_{\odot}c^2$ in 10^{10} years; see

Chapter I) with the cumulative mass loss by a typical spiral galaxy. This quantity was shown to be of order $\sim 10^9\ M_\odot$ for the bulge component of our Galaxy, so stellar mass loss can indeed fuel observed Seyfert activity, even if energy is produced only at 'nuclear' efficiences ($\sim 0{\cdot}7$ %).

V DETAILED THEORY OF THE FORMATION AND STRUCTURE OF THE DISC

This chapter deals at an order of magnitude level with the very complex problem of the structure of the gaseous disc formed by inflowing material lost from the galaxy's bulge-component stars. The main aims are two-fold: First, to isolate the most important physical processes and obtain results that are at least qualitatively correct; and secondly, to provide convincing theoretical arguments that the formation of an active nucleus is indeed an almost inevitable process. The theory, even at this order of magnitude level, becomes quite involved in some sections (due partly to the number of separate cases requiring discussion, and partly to the inherent complexity of the problem), and for this reason, by way of introduction, it was thought worthwhile to include the following brief summary of the problem.

Whatever the nature of observed nuclear activity, be it due to a spinar or a black hole, one of the major difficulties that must be overcome is how to explain inward flows of matter in excess of $\sim 0{\cdot}1\ M_\odot\ y^{-1}$. This is the case whichever model is assumed. For example, if nuclear activity is produced by a period of rapid accretion onto a black hole, the luminosity may be expressed in the form

$$L \sim \eta\, \dot{M}_{acc}\, c^2$$

where η is the efficiency with which matter is converted into energy by the accretion process and $\dot{M}_{acc}$ is the accretion rate. For disc accretion, Thorne (1974) has shown that η is unlikely to much exceed $\sim 30\ \%$; so we have

$$L \sim 1{\cdot}7 \times 10^{39} \left(\frac{\eta}{0{\cdot}3}\right)\left(\frac{\dot{M}_{acc}}{1 M_\odot y^{-1}}\right) \quad W \qquad (13)$$

The luminosity of a typical Seyfert galaxy is $\sim 10^{38}$ W, so the black hole model must be able to explain central fluxes of order $0{\cdot}1\ M_\odot\ y^{-1}$. If the same process is responsible for the luminous quasar phenomenon (with L up to $\sim 10^{41}$ W), fluxes as high as $\sim 50\ M_\odot\ y^{-1}$ must occur. On the other hand, if nuclear activity is thought to be caused by the evolution of a massive object such as a spinar (e.g. Ozernoy & Usov, 1971, 1973), it is necessary that the body be formed on a timescale less than that on which it evolves. A spinar's expected evolutionary timescale is given approximately by $\tau_{evol} \sim 10^6$ y (depending on the detailed model, its mass and luminosity), so provided that $M \gtrsim 10^5\ M_\odot$, central fluxes $\gtrsim 0{\cdot}1\ M_\odot\ y^{-1}$ are again indicated.

In the present theory, matter is transported inwards by the action of viscous friction (Lynden-Bell & Pringle, 1974). Some results from viscous disc theory are reviewed in Section V-2, and it is noted that the expected central flux is proportional to $\nu \mathcal{M}_d$ where ν is the kinematic viscosity and $\mathcal{M}_d$ the disc's mass. Since continuing infall from stellar mass loss leads to a slow increase in $\mathcal{M}_d$, if ν was assumed to be actually independent of time an eventual steady-state would be reached in which the central flux equalled the net infall rate. However stellar mass loss from a typical nuclear star cluster of mass $M_n \sim 10^{10}\ M_\odot$ occurs at present epochs (i.e. $t \sim 10^{10}$ y) at a rate only on the order of $10^{-2}\ M_\odot\ y^{-1}$, too small to explain observed activity. Thus, evolution at constant ν does not explain Seyfert activity, and if stellar mass loss is the source of fuel for observed activity ν must depend rather sensitively on the disc's mass. In order to achieve high enough central fluxes for short periods (so that the average central flux still equals the net infall rate), the form of the mass-dependence must be such that ν is small while $\mathcal{M}_d < M_{crit}$, and then, at

$\mathcal{M}_d \sim M_{crit}$, it must suddenly increase to a much larger value. Since ν is itself proportional to the square of the mean turbulent velocity of the gas (Section V-2), an obvious candidate for causing such a mass-dependence is the possibility of a star-forming gravitational instability in which a cold dense disc (characterised by $\bar{c} \sim 1$ km s^{-1}) is rapidly converted into a system in which $\bar{c} \sim 20$ km s^{-1}, representative of the velocity dispersion in newly-formed H II regions. Such an event would increase the central flux by more than two orders of magnitude.

A major aim of this chapter is to investigate, by means of a detailed study of the disc's structure, whether it is indeed possible for the sequence of events outlined above to occur and lead to a period of nuclear activity. Because it was impractical to consider a general model capable of extension to a wide variety of different galaxies, it was thought best to base the discussion on a particular model of the central regions of our own Galaxy. In this way, although direct extension of the results to elliptical galaxies and the radio-galaxy phenomenon is not possible, the detailed results of the theory ought to apply to most typical spiral galaxies.

V-1 Formation and surface density of the disc

This section deals briefly with the assumed star distribution in the nuclear star cluster, and presents a calculation of the equilibrium surface densities of two representative discs formed by stellar mass loss within such a cluster. The assumed star distribution is based on the model of our Galactic Centre constructed by Sanders & Lowinger (1972), who found that a wide variety of observational data could be fitted by a flattened spheroidal model in which the star density decreased outwards (from a central core of radius $\sim 0{\cdot}5$ pc) as a power-law roughly proportional to $r^{-1{\cdot}8}$. The total stellar mass contained within a radius $R_n \sim 1$ kpc was $M_n \sim 10^{10}\,M_\odot$.

In this investigation the star density is approximated by a spherically symmetric model defined by

$$\rho_*(r) = \frac{\mathcal{M}_n}{4\pi R_n} \cdot \frac{1}{r^2} \qquad 0 \leq r \leq R_n \qquad (14)$$

This distribution has two advantages. First, from an observational viewpoint it closely approximates the Sanders-Lowinger model of our Galaxy, which means that results of the present theory ought to apply in particular to our own Galaxy; and secondly, from a theoretical viewpoint it has the advantage of resembling at large radii the mass distribution of an isothermal sphere. However the lack of a nearly constant density 'core', and the divergence of $\rho_*(r)$ at the origin, means that results referring specifically to small radii ($r \lesssim 1$ pc, say) are unlikely to have much physical validity.

The projected star distribution is given by

$$\sigma_*(\varpi) = 2\int_0^\infty \rho_*(\varpi, z)\,dz$$

which, by use of (14), reduces to

$$\sigma_*(\varpi) = \frac{\mathfrak{M}_n}{2\pi R_n}\,\frac{1}{\varpi}\,\cos^{-1}\left(\varpi/R_n\right) \qquad (15)$$

where ϖ is the cylindrical radial coordinate, defined by $\varpi^2 = r^2 - z^2$, and z is measured perpendicular to the disc.

An important parameter in the theory is the assumed distribution of angular momentum. In order to model the effects of net angular momentum in the assumed spherically symmetric cluster, the distribution (14) is divided into a number of concentric cylindrical shells (aligned with axes parallel to the z-axis), such that each has a net rotational velocity $V_{rot}(\varpi)$. Observations of the central region of our Galaxy do not yet allow a choice to be made as to the most representative functional form of V_{rot}, so in order to try and bracket the likely range of possibilities two particular forms were considered. In the first, (Case 'A') it was assumed that the net rotation was independent of ϖ, whereas in the second, (Case 'B') V_{rot} was assumed to correspond to uniform rotation; i.e.

CASE A: $V_{rot}(\varpi) = V_n$

CASE B: $V_{rot}(\varpi) = V_n\left(\frac{\varpi}{R_n}\right)$

where V_n is defined to be the net rotational velocity at $\varpi = R_n$.

In order to calculate the equilibrium disc surface density, its formation was approximated by a two-stage process. First the gas lost from stars was assumed to flow in towards mid-plane, where it formed a 'projected disc' with density distribution proportional to the projected star density; and secondly, this disc was then assumed to contract to its final form in such a way that each annular element conserved both mass and angular momentum. The surface density of the projected disc is

$$\sigma_p(\varpi) = \frac{\mathcal{M}_d}{2\pi R_n} \frac{1}{\varpi} \cos^{-1}(\varpi/R_n) \qquad (16)$$

and if contraction from ϖ to ϖ' occurs, conservation of mass implies

$$\sigma(\varpi') = \frac{\varpi}{\varpi'} \frac{d\varpi}{d\varpi'} \sigma_p(\varpi) \qquad (17)$$

Detailed conservation of angular momentum gives

$$\varpi V_{rot}(\varpi) = \varpi' V_c(\varpi') \qquad (18)$$

where $V_c(\varpi)$ is the circular velocity at radius ϖ. In the present model the disc mass within ϖ is always negligible compared to the stellar mass contained with the same radius, so $V_c(\varpi)$ is very nearly equal to the circular velocity due to the stars alone. With this approximation (14) implies $V_c(\varpi) = (GM_n/R_n)^{\frac{1}{2}} = V_o$, say. (18) thus reduces to

$$\varpi V_{rot}(\varpi) = \varpi' V_o$$

The disc radius, R_d, is defined by

$$R_d = R_n V_n / V_o$$

so using equations (16) - (18) the expressions for the final equilibrium surface densities in Cases A and B become:

$$\sigma_A(\varpi) = \frac{\mathfrak{M}_d}{2\pi R_d} \cdot \frac{1}{\varpi} \cdot \cos^{-1}\left(\varpi/R_d\right) \tag{19-A}$$

$$\sigma_B(\varpi) = \frac{\mathfrak{M}_d}{4\pi} \cdot \frac{1}{R_d^{1/2}\,\varpi^{3/2}} \cdot \cos^{-1}\left(\sqrt{\varpi/R_d}\right) \tag{19-B}$$

V-2 Viscous disc theory

In the theory, matter is stored in a dense gaseous disc, and then, after the onset of a star-forming gravitational instability, transported into the origin by the action of viscosity. It is therefore necessary to discuss some results from viscous disc theory (Lynden-Bell & Pringle, 1974), and in particular estimate the order of magnitude of the kinematic viscosity ν . This may be written (e.g. Lynden-Bell & Pringle, 1974) in the form

$$\nu \sim \frac{1}{3} \bar{c} \, l \qquad (20)$$

where $\bar{c}$ is the mean random velocity of an 'element' of the disc (i.e. a cloud, molecule, or turbulent eddy), and l is an effective mean free path. In the case of molecular viscosity, l is the same as the molecular mean free path, but in the case of viscosity due to cloud-cloud collisions l may be substantially different from the collision mean free path. A realistic determination of ν during the two principal phases of the disc's evolution is an extremely difficult and complex problem, as it depends to a great extent on the detailed physical model of the disc that is assumed. A discussion of this question is given in the Appendix, and it is shown that the kinematic viscosity during both phases of the disc's evolution may be estimated approximately by equating $\bar{c}$ to the sound speed and letting l equal the mean amplitude of the epicyclic motions in the plane; i.e.

$$\bar{c} \sim V_{sound}$$

and

$$l \sim \bar{c}/\kappa$$

where κ is the epicyclic frequency defined in terms of the usual Oort constants by $\kappa^2 = 4B(B-A)$. The kinematic viscosity is therefore given approximately by

$$\nu \sim \frac{1}{3}\, \bar{c}^2/\kappa$$

where it should be noted (see the Appendix) that the eaxact numerical coefficient (which depends on the detailed model) is possibly uncertain by a factor as large as 5. Using the definitions

$$A = \frac{1}{2}\left(\frac{V_c}{\varpi} - \frac{dV_c}{d\varpi}\right)$$

and

$$B = -\frac{1}{2}\left(\frac{V_c}{\varpi} + \frac{dV_c}{d\varpi}\right)$$

therefore gives

$$\nu \sim \frac{1}{3\sqrt{2}}\ \bar{c}^2\ \frac{\varpi}{V_c(\varpi)}\left(1 + \frac{\varpi}{V_c}\frac{dV_c}{d\varpi}\right)^{-1/2} \qquad (21)$$

If the disc is in a steady state with no central couple, the inward flux, F, of matter across any radius ϖ is constant. This flux is given (Lynden-Bell & Pringle, 1974) by

$$F = g/h$$

where h is the specific angular momentum at radius ϖ and g is the couple exerted by the material inside ϖ on that outside ϖ. Since $g = 2\pi\varpi^2\nu.\sigma.2A$ and $h = \varpi V_c(\varpi)$, this can be re-written in the form

$$F = 2\pi\nu\sigma\left(1 - \frac{\varpi}{V_c}\frac{dV_c}{d\varpi}\right) \qquad (22)$$

In the present model $V_c = V_o$ and $\nu \propto \varpi$, so the equilibrium surface density is $\sigma_{eq} \propto \varpi^{-1}$

An important quantity to establish is the timescale on which viscous effects might be important. The viscous timescale (defined here as the time required for material at radius ϖ to reach the origin in a steady state disc) is given by

$$\tau_{visc} \sim \varpi / u_{\varpi}$$

where u_{ϖ} is the inward drift velocity of material at ϖ, related to F by

$$F = 2\pi \varpi u_{\varpi}$$

Combining these expressions allows the viscous timescale to be written as

$$\tau_{visc} \sim \varpi^2 / \nu \left(1 - \frac{\varpi}{V_c} \frac{dV_c}{d\varpi} \right)$$

$$\text{i.e.} \quad \tau_{visc} \sim 3\sqrt{2} \; \frac{\varpi V_c(\varpi)}{\bar{c}^2} \; \frac{\left(1 + \frac{\varpi}{V_c} \frac{dV_c}{d\varpi} \right)^{1/2}}{\left(1 - \frac{\varpi}{V_c} \frac{dV_c}{d\varpi} \right)} \qquad (23)$$

With the assumed star distribution, and adopting $M_n \sim 10^{10}\ M_\odot$, $R_n \sim 1$ kpc and $V_c(\varpi) = V_o \sim 207{\cdot}4$ km s^{-1}, this reduces to

$$\tau_{visc} \sim 8{\cdot}6 \times 10^6 \left(\frac{\varpi}{1\,\mathrm{pc}} \right) \left(\frac{10\ \mathrm{km\,s^{-1}}}{\bar{c}} \right)^2 \ \mathrm{y} \qquad (24)$$

Inspection of this expression shows that at radii $\varpi \gtrsim 1$ pc viscous evolution may be neglected on timescales $\lesssim 10^9$ y provided that $\bar{c}$ is small ($\lesssim 1$ km s^{-1}, say). Close to the centre ($\varpi \lesssim 1$ pc) it is possible that viscous evolution <u>might</u> occur even if $\bar{c}$ is small; but because the assumed star distribution is unphysical in this region, an improved calculation of the central disc density allowing for possible viscous evolution was not thought worthwhile. In a more realistic model the central star density is likely to approach a constant value, and in this case (provided that the nucleus does not already contain a massive collapsed remnant such as a black hole) the central parts of the disc ($\varpi \lesssim 1$ pc) are likely to be characterised by nearly uniform rotation. Although deviations from exactly uniform rotation may still occur very close to the centre ($\varpi \ll 1$ pc) where the disc mass within ϖ becomes comparable with the stellar mass within the same radius, inspection of (23) shows that the viscous timescale throughout most of this central region is likely to be much longer than that given by (24). In the discussion that follows, complications due to possible viscous evolution in the very central region are ignored, and it is assumed that viscous evolution prior to the onset of gravitational instability can be neglected throughout the disc provided that $\bar{c}$ is sufficiently small.

V-3 Critical disc density

The stability of self-gravitating discs has been discussed by a number of authors, in particular Goldreich & Lynden-Bell (1965a,b), from whose work most of the results used here will be drawn. They showed that significant growth of small perturbations would occur whenever the relation

$$\frac{\pi G \bar{\rho}}{\kappa^2} \gtrsim 1 \qquad (25)$$

(accurate to within a factor of order 2) was satisfied. Here κ is again the epicyclic frequency, and $\bar{\rho}$ is defined by

$$\bar{\rho} = \int \rho^2(\varpi, z)\, dz \Big/ \int \rho(\varpi, z)\, dz$$

Adopting an isothermal approximation to the disc's structure, in which the r.m.s. velocity dispersion V_s is assumed to be independent of z, and neglecting the gravitational influence of the star distribution on the disc's z-structure, it may be shown that the disc density is

$$\frac{\rho(\varpi, z)}{\rho(\varpi, 0)} = 1 \Big/ \cosh^2(\omega \rho(\varpi, 0) z) \qquad (26)$$

where the weight factor w (following Goldreich & Lynden-Bell's notation) is defined by

$$\omega^2 = 6\pi G \Big/ \rho(\varpi, 0) V_s^2$$

The disc width, here defined to be

$$H(\varpi) = \sigma(\varpi) \Big/ \rho(\varpi, 0) \qquad (27)$$

is therefore

$$H(\Theta) = 2\int_0^\infty \frac{dz}{\cosh^2(\omega\rho(\Theta,0)z)} = 2/\omega\rho(\Theta,0)$$

$$\text{i.e} \quad H(\Theta) = \frac{2}{3\pi}\frac{V_s^2}{G\sigma(\Theta)} \tag{28}$$

$\bar{\rho}$ may be shown to equal $\frac{2}{3}\rho(\Theta,0)$, so (25) reduces to

$$\frac{\pi G\sigma}{\kappa V_s} \gtrsim 1$$

which implies the critical surface density σ_{crit} is given by

$$\sigma_{crit} \sim \frac{\kappa V_s}{\pi G} \tag{29}$$

This criterion is identical to that used by Goldreich & Ward (1973) in their discussion of the fragmentation of a dust disc during an early stage of solar system evolution.

Inserting this value of σ into the formula for the central flux in a steady state viscous disc (equation (22)) allows an estimate of F to be made:

$$F_{crit} \sim 2\pi\nu\sigma_{crit}$$

$$\Rightarrow F_{crit} \sim \frac{2}{3}\frac{\bar{c}^2 V_s}{G}$$

$$\text{i.e.} \quad F_{crit} \sim 1{\cdot}6\times10^{-4}\left(\frac{\bar{c}}{1\,\mathrm{km\,s^{-1}}}\right)^2\left(\frac{V_s}{1\,\mathrm{km\,s^{-1}}}\right) M_\odot\, y^{-1} \tag{30}$$

Thus, provided that the disc is gravitationally stable ($\sigma < \sigma_{crit}$) and that it can remain sufficiently cool during this phase of its evolution ($\bar{c} \lesssim 1$ km s^{-1}, say), the inward viscous flux will be very much less than the stellar mass loss rate and the disc will slowly grow in mass. Eventually a star-forming gravitational instability will occur, and $\bar{c}$ may be expected to increase to a value similar to that observed in many H II regions ($\bar{c} \sim 20$ km s^{-1}, say). Substituting this value into (30) then implies inward transport at a rate of order $0{\cdot}05\ M_{\odot}\ y^{-1}$, which is sufficient to explain typical Seyfert activity by either accretion onto a massive black hole (cf. equation (13)) or by the formation and evolution of some other kind of massive object.

V-4 Volume density

In order to determine the heating and cooling rates in the disc (see Sections V-5 and V-6 below), it is necessary first to calculate the distribution of disc material perpendicular to the plane of symmetry. Following Spitzer (1968; p.180) the equation of hydrostatic balance in the z-direction is written in the form

$$\frac{1}{3\rho}\frac{\partial}{\partial z}\left(\rho V_s^2\right) = -\frac{\partial \phi}{\partial z} \qquad (31)$$

where V_s is the total r.m.s. gas velocity (including possible turbulent motions), so that the gas pressure is

$$P_{gas} = \frac{1}{3}\rho V_s^2$$

The effects of cosmic rays and magnetic fields have been ignored in this preliminary investigation, although their effective pressures may be represented in any of the following formulae by simply writing $V_s(1+\alpha+\beta)^{\frac{1}{2}}$ everywhere that V_s occurs. (α and β - assumed independent of z - are the ratios respectively of cosmic ray pressure and magnetic field pressure to total gas pressure).

Poisson's equation for the disc close to its plane of symmetry may be approximated by

$$\nabla^2\phi_{gas} \sim \frac{\partial^2\phi_{gas}}{\partial z^2} \sim 4\pi G\rho \quad .$$

so in the 'isothermal' approximation (in which V_s does not depend on z, and radial pressure gradients are ignored) (31) may be re-written

$$\frac{1}{3}V_s^2\frac{\partial}{\partial z}\left(\frac{1}{\rho}\frac{\partial\rho}{\partial z}\right) = -4\pi G\rho - \frac{\partial^2\phi_*}{\partial z^2} \qquad (32)$$

Although $\frac{\partial^2 \phi_*}{\partial z^2}$ is in principle a known function (since it depends only on the assumed star distribution), the general solution of (32) is not easily obtained. However limiting solutions may be determined in two extreme cases: (1) self-gravity negligible; and (2) self-gravity dominant. In the first of these (32) becomes

$$\frac{\partial}{\partial z}\left(\frac{1}{\rho}\frac{\partial \rho}{\partial z}\right) = -\frac{3}{v_s^2 V_0^2}\frac{\partial^2 \phi_*}{\partial z^2} \tag{33}$$

where the dimensionless velocity dispersion v_s is defined by

$$v_s \equiv V_s / V_0 \tag{34}$$

The potential, $\phi_*(r)$, is given for the assumed star distribution (14) by

$$\phi_*(r) = V_0^2 \left[\ln(r/R_n) - 1 \right]$$

where it has been assumed that $\phi_*(r \geq R_n)$ is given by $-V_0^2.(R_n/r)$.
(33) thus reduces to

$$\frac{\rho(\varpi, z)}{\rho(\varpi, 0)} = \left(1 + z^2/\varpi^2\right)^{-3/2v_s^2}$$

Provided that v_s is small ($\lesssim 0{\cdot}1$, say), this may be closely approximated for small z by

$$\frac{\rho(\varpi, z)}{\rho(\varpi, 0)} \sim \exp\left(-\frac{3}{2v_s^2}\frac{z^2}{\varpi^2}\right) \tag{35}$$

The disc width, defined by $H(\varpi) = \sigma(\varpi)/\rho(\varpi, 0)$, is therefore given by

$$H(\varpi) = \sqrt{\frac{2\pi}{3}}\, v_s\, \varpi \tag{36}$$

In the opposite extreme, that in which self-gravity is dominant, it may be shown (e.g. Goldreich & Lynden-Bell, 1965a) that the solution is given by (26); i.e.

$$\frac{\rho(\varpi,z)}{\rho(\varpi,0)} = 1 \Big/ \cosh^2\left(\omega\, \rho(\varpi,0)\, z\right) \qquad (26)$$

where

$$\omega^2 = 6\pi G \Big/ \rho(\varpi,0)\, V_s^2$$

In this limit the disc width depends on both V_s and σ, and is

$$H(\varpi) = \frac{2}{3\pi}\,\frac{V_s^2}{G\sigma} \qquad (28)$$

In general the disc width will be a little smaller than that given by either (36) or (28), but it may be verified that discs which are close to instability are well described by the approximation in which self-gravity is dominant (as it must be for discs which are gravitationally unstable). When gravitational instability does occur stars may be expected to form, and if some of these are sufficiently massive to drive strong mass motions ('turbulence') V_s may well increase by a large factor. In this event, and also if $\sigma \ll \sigma_{crit}$, the disc structure will be better described by the approximation (35).

For the purpose of the order of magnitude estimates arrived at in the next sections it is assumed that the disc can be approximated by an 'equivalent disc', defined here to be one having the same surface density as the actual disc, width $H(\varpi)$ and constant density equal to $\rho(\varpi,0)$ in the z-direction. The density of this disc (or the mid-plane density of the actual disc) depends not only on the particular rotation law under discussion (i.e. whether Case A or Case B), but also on whether or not the disc density is such that self-gravity can be ignored. There are therefore four possibilities to consider:

Case A1 : V_{rot} = const; self-gravity negligible.

Case A2 : V_{rot} = const; self-gravity dominant.

Case B1 : $V_{rot} \propto \varpi$; self-gravity negligible.

Case B2 : $V_{rot} \propto \varpi$; self-gravity dominant.

Substituting for $\sigma(\varpi)$ and $H(\varpi)$, the mid-plane density for each possibility becomes

$$\rho_{A1}(\varpi,0) \sim \frac{\mathcal{M}_d}{(2\pi)^{3/2}} (3)^{1/2} \frac{1}{v_s} \frac{1}{R_d} \frac{1}{\varpi^2} \cos^{-1}(\varpi/R_d) \qquad (37\text{-}A1)$$

$$\rho_{A2}(\varpi,0) \sim \frac{3}{8\pi} \frac{\mathcal{M}_d^2}{\mathcal{M}_n} \frac{1}{v_s^2} \cdot \frac{R_n}{R_d^2} \frac{1}{\varpi^2} \left(\cos^{-1}(\varpi/R_d)\right)^2 \qquad (37\text{-}A2)$$

$$\rho_{B1}(\varpi,0) \sim \frac{\mathcal{M}_d}{(2\pi)^{3/2}} \left(\frac{3}{4}\right)^{1/2} \frac{1}{v_s} \frac{1}{R_d^{1/2}} \frac{1}{\varpi^{5/2}} \cos^{-1}\left(\sqrt{\varpi/R_d}\right) \qquad (37\text{-}B1)$$

$$\rho_{B2}(\varpi,0) \sim \frac{3}{32\pi} \frac{\mathcal{M}_d^2}{\mathcal{M}_n} \frac{1}{v_s^2} \frac{R_n}{R_d} \frac{1}{\varpi^3} \left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^2 \qquad (37\text{-}B2)$$

If random gas velocities are assumed to be on the order of the sound speed (since supersonic motions are quickly dissipated in a dense gaseous medium) the r.m.s. gas velocity V_s is related to the sound speed by the approximate equality $V_s = \sqrt{2}\, V_{sound}$. To allow for the possibility that $\bar{c}$ will not be exactly equal to V_{sound}, the relation ought properly to be expressed in the form

$$V_s = \xi \sqrt{2}\, V_{sound} \tag{38}$$

where ξ is a constant of order unity, but in the interests of clarity this paramater will be ignored.

The sound speed is defined by

$$V_{sound} = \left(3kT/\bar{m}\right)^{1/2} \tag{39}$$

and depends not only on the temperature but also on the mean particle mass $\bar{m}$. In the numerical estimates obtained below it is assumed that $\bar{m} = 1{\cdot}3\ m_H$ for an atomic disc, and $2{\cdot}5\ m_H$ when the disc is predominantly in molecular form. m_H is taken to be $1{\cdot}7 \times 10^{-27}$ kg. With these parameters, and assuming $V_o = 207{\cdot}4$ km s^{-1} (appropriate for a star cluster with $M_n = 10^{10}\ M_\odot$ and $R_n = 1$ kpc), we obtain

$$\left.\begin{aligned} v_{s,\,atomic} &\sim 9{\cdot}3 \times 10^{-4}\ T^{1/2} \\ v_{s,\,molecular} &\sim 6{\cdot}7 \times 10^{-4}\ T^{1/2} \end{aligned}\right\} \tag{40}$$

A useful quantity that is required at a later stage of the argument is the particle number density, n, defined by

$$n(\varpi) = \rho(\varpi, 0)/\bar{m}$$

For the cases under discussion this is

$$n_{A1} = C_{A1} \cos^{-1}\left(\mathcal{D}/R_d\right)\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right) T^{-1/2} \left(\frac{500\,\mathrm{pc}}{R_d}\right)\left(\frac{100\,\mathrm{pc}}{\mathcal{D}}\right)^2 \quad m^{-3} \qquad (41\text{-}A1)$$

$$n_{A2} = C_{A2} \left(\cos^{-1}\left(\mathcal{D}/R_d\right)\right)^2\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)^2 T^{-1} \left(\frac{500\,\mathrm{pc}}{R_d}\right)\left(\frac{100\,\mathrm{pc}}{\mathcal{D}}\right)^2 \quad m^{-3} \qquad (41\text{-}A2)$$

$$n_{B1} = C_{B1} \cos^{-1}\left(\sqrt{\mathcal{D}/R_d}\right)\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right) T^{-1/2} \left(\frac{500\,\mathrm{pc}}{R_d}\right)\left(\frac{100\,\mathrm{pc}}{\mathcal{D}}\right)^{5/2} \quad m^{-3} \qquad (41\text{-}B1)$$

$$n_{B2} = C_{B2} \left(\cos^{-1}\left(\sqrt{\mathcal{D}/R_d}\right)\right)^2\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)^2 T^{-1} \left(\frac{500\,\mathrm{pc}}{R_d}\right)\left(\frac{100\,\mathrm{pc}}{\mathcal{D}}\right)^3 \quad m^{-3} \qquad (41\text{-}B2)$$

where the constants C are given by

$$C_{A1} = \begin{cases} 7{\cdot}25 \times 10^9 & \text{for an atomic disc} \\ 5{\cdot}23 \times 10^9 & \text{for a molecular disc} \end{cases}$$

$$C_{A2} = 1{\cdot}69 \times 10^{10} \quad \text{for both atomic and molecular discs}$$

$$C_{B1} = \begin{cases} 8{\cdot}10 \times 10^9 & \text{for an atomic disc} \\ 5{\cdot}85 \times 10^9 & \text{for a molecular disc} \end{cases}$$

$$C_{B2} = 2{\cdot}10 \times 10^{10} \quad \text{for both atomic and molecular discs.}$$

V-5 Heating processes

Four sources of heat have been considered as of possible importance to the energetics of the disc, namely

(1) Heating by infalling material, Γ_1

(2) Heating by stellar radiation, Γ_2

(3) Heating by supernovae in the disc, Γ_3

(4) Heating by stars passing through the disc, Γ_4

(1) Heating by infalling material

In practice infalling gas may be expected to preferentially heat the upper layers of the disc, but since both the discussion leading to the instability criterion (Section V-3) and also that leading to the equilibrium density distribution has been based on the 'isothermal' approximation, this complicating factor has been ignored. In addition, in order to obtain rough estimates of the heating and cooling rates, the disc's structure is further simplified, and it is assumed that the actual disc can be represented for these purposes by an 'equivalent disc', defined to have the same surface density as the actual disc, but constant density in the z-direction equal to $\rho(\varpi, 0)$. The width of the equivalent disc is $H(\varpi)$.

In order to calculate the heating effect due to infalling material, it is assumed that material entering the disc at deposits its energy uniformly throughout the disc's z-extent. In this approximation the volumetric heating rate due to infall is

$$\Gamma_1(\varpi) \sim \epsilon(\varpi) \frac{\partial \rho}{\partial t} \qquad (42)$$

where $\epsilon(\varpi)$ is the energy deposited by unit mass entering the disc at ϖ, and where it is also assumed that material becoming part of the disc at radii other than ϖ does not contribute significantly to $\Gamma_1'(\varpi)$. If changes with time of $H(\varpi)$ are neglected, (42) becomes

$$\Gamma_1 \sim \frac{\epsilon(\varpi)}{H(\varpi)} \frac{\partial \sigma}{\partial t} \qquad (43)$$

and if σ is assumed to change only through infall (i.e. viscous evolution is neglected),

$$\frac{\partial \sigma}{\partial t} = \frac{\sigma(\varpi)}{\mathcal{M}_d} \frac{d\mathcal{M}_d}{dt} = \alpha_* \mathcal{M}_n \frac{\sigma(\varpi)}{\mathcal{M}_d}$$

Thus (43) becomes

$$\Gamma_1(\varpi) \sim \epsilon(\varpi) \frac{1}{H(\varpi)} \alpha_* \mathcal{M}_n \frac{\sigma(\varpi)}{\mathcal{M}_d} \qquad (44)$$

It is now necessary to estimate $\epsilon(\varpi)$, the energy deposited by unit mass entering the disc at radius ϖ. It is convenient to divide this into two terms: a gravitational term representing the gravitational energy that is lost by material becoming part of the disc; and a thermal term representing its ititial thermal energy.

$$\text{i.e.} \qquad \epsilon(\varpi) = \epsilon_{grav}(\varpi) + \epsilon_{th}(\varpi)$$

On general grounds it may be expected that the gravitational part of ϵ will be of order V_o^2, although the appropriate numerical factor, being larger for cases in which the contraction of the 'projected disc' to its final form is largest, is obviously model-dependent. The thermal part of ϵ is more difficult to estimate, as it depends not only on what fraction of the gas' internal energy is radiated before entering the disc, but also on how efficiently supernova heat the gas in the first place. If supernova heating is very efficient, as it is in situations leading to thermally steady flows (cf. Mathews & Baker, 1971), the temperature of the infalling gas might be as high as that normally associated with a hot outflowing galactic wind; i.e. $T \sim 2 \times 10^7$ K. In this case, assuming the gas to be pure ionised hydrogen, ϵ_{th} is given by $\epsilon'_{th} \sim 3kT/m_H$ with $T \sim 2 \times 10^7$ K. In order to allow for the possibility that the infalling material is cold, or has managed to radiate a significant fraction of its initial thermal energy before entering the disc, ϵ_{th} is written here in the form $\epsilon_{th} \sim f_1 \epsilon'_{th}$, where the fraction f_1 $(0 \leq f_1 \leq 1)$ remains undetermined. Adopting $T \sim 2 \times 10^7$ K and taking the previously adopted value for V_0 ($207{\cdot}4$ km s^{-1}) thus implies

$$\epsilon(\varpi) = a(\varpi) V_0^2 \qquad (45)$$

where the factor a, expected to lie in the approximate range $1 \lesssim a \lesssim 10$, is defined by

$$a(\varpi) \sim \left(10 f_1 + \epsilon_{grav}/V_0^2\right) \qquad (46)$$

Γ_1 is therefore

$$\Gamma_1 \sim a V_0^2 \frac{\alpha_* M_1}{M_d} \rho(\varpi, 0) \qquad (47)$$

(2) <u>Heating by stars within the disc</u>

The heating effect due to stellar radiation is easily obtained from the assumed star distribution by multiplying the star density by a mean mass-luminosity ratio and a small factor f_2 to represent the fraction of the incident stellar radiation that is actually converted into random gas motions. The volumetric heating rate due to stars in the disc is thus

$$\Gamma_2 \sim f_2 \; \overline{L}_* \; n_*$$

where n_* is the space density of stars, given by $n_* = \rho_* / \bar{m}_*$. This implies

$$\Gamma_2 \sim \frac{M_n}{4\pi R_n} \frac{1}{\varpi^2} \left(\frac{\overline{L}_*}{\bar{m}_*} \right) \cdot f_2 \qquad (48)$$

where the star distribution (14) has been assumed.

(3) Heating by supernovae within the disc

The calculation of Γ_1 above allowed for the possibility that supernovae at large distances from the disc might heat the infalling material, but it is still of interest to enquire as to the heating by supernovae located actually within the disc. If supernovae occur in the nuclear star cluster with a mean interval of τ (e.g. τ years per mass M_n of stars) the supernova rate per unit volume is

$$S = \rho_* / M_n \tau$$

$$\Rightarrow \quad S = \frac{1}{4\pi R_n} \frac{1}{\varpi^2} \frac{1}{\tau} \qquad \text{supernovae } m^{-3}\, s^{-1}$$

Thus, assuming that each supernova deposits energy ϵ_{SN} into the disc, the supernova heating rate is

$$\Gamma_3 \sim \epsilon_{SN} \frac{1}{4\pi R_n} \frac{1}{\varpi^2} \frac{1}{\tau} \qquad (49)$$

(4) Heading by stars passing through the disc

This heating effect is estimated approximately in the following way. A star passing through the disc gives gas distance b away an impulse

$$\delta v \sim \frac{2 G m_*}{b V_*}$$

In unit time the star travels a distance V_* through the disc, so if the gas density is ρ the energy deposited per unit time into a shell of radius b and thickness db is

$$d\dot{E} = \frac{1}{2} \rho (\delta v)^2 \, 2\pi b V_* \, db$$

$$\Rightarrow \quad d\dot{E} = \frac{4\pi G^2 m_*^2}{V_*} \rho \frac{db}{b}$$

The volumetric heating rate due to stars passing through the disc is therefore $\Gamma_4 \sim n_* \dot{E}$

$$\text{i.e } \Gamma_4 \sim G \bar{m}_* \frac{V_0^2}{V_*} \frac{\rho}{\varpi^2} \ln\left(\frac{b_{max}}{b_{min}}\right) \qquad (50)$$

where the relation $V_0^2 = GM_n/R_n$ has been used, and b_{max} and b_{min} are the maximum and minimum impact parameters respectively. Assuming $V_* \sim V_0$ and $\ln(b_{max}/b_{min}) \sim 10$, thus implies

$$\Gamma_4 \sim 10 \, G \, \bar{m}_* V_0 \frac{\rho(\varpi,0)}{\varpi^2} \qquad (51)$$

The total heating rate per unit volume is therefore

$$\Gamma = \sum_{i=1}^{4} \Gamma_i$$

i.e.

$$\Gamma = a V_0^2 \alpha_* M_n \frac{\rho(\varpi,0)}{M_d}$$

$$+ f_2 \frac{M_n}{4\pi R_n} \frac{1}{\varpi^2} \left(\frac{\bar{L}_*}{\bar{m}_*} \right)$$

$$+ \epsilon_{SN} \frac{1}{4\pi R_n} \frac{1}{\varpi^2} \frac{1}{\tau}$$

$$+ 10\, G\, \bar{m}_* V_0 \frac{\rho(\varpi,0)}{\varpi^2} \qquad (52)$$

It is now shown that the total heating is dominated by infalling material; i.e.

$$\Gamma \sim \Gamma_1 \sim a V_0^2 \alpha_* M_n \frac{\bar{m}\, n(\varpi)}{M_d}$$

Substituting $n(\varpi)$ (equations (41)) into the expression for Γ_1 gives (for $M_n \sim 10^{10}\, M_\odot$ and $R_n \sim 1$ kpc):

$$\left.\begin{aligned} \Gamma_{1,A1,\,atomic} &\sim \frac{2\cdot18\times10^{-23}\, a \cos^{-1}(\varpi/R_d)}{T^{1/2}} \left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{500\,pc}{R_d}\right)\left(\frac{100\,pc}{\varpi}\right)^2 \;\; W m^{-3} \\ \Gamma_{1,A1,\,molecular} &\sim \frac{3\cdot03\times10^{-23}\, a \cos^{-1}(\varpi/R_d)}{T^{1/2}} \left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{500\,pc}{R_d}\right)\left(\frac{100\,pc}{\varpi}\right)^2 \;\; W m^{-3} \end{aligned}\right\} \quad (53\text{-}A1)$$

$$\left.\begin{aligned} \Gamma_{1,A2,\,atomic} &\sim \frac{5\cdot10\times10^{-23}\, a \left(\cos^{-1}(\varpi/R_d)\right)^2}{T} \left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{M_d}{10^{7} M_\odot}\right)\left(\frac{500\,pc}{R_d}\right)^2\left(\frac{100\,pc}{\varpi}\right)^2 \;\; W m^{-3} \\ \Gamma_{1,A2,\,molecular} &\sim \frac{9\cdot79\times10^{-23}\, a \left(\cos^{-1}(\varpi/R_d)\right)^2}{T} \left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{M_d}{10^{7} M_\odot}\right)\left(\frac{500\,pc}{R_d}\right)\left(\frac{100\,pc}{\varpi}\right)^2 \;\; W m^{-3} \end{aligned}\right\} \quad (53\text{-}A2)$$

$$\left.\begin{aligned}\Gamma_{1,B1,\text{atomic}} &\sim \frac{2\cdot 44\times 10^{-23}\, a\, \cos^{-1}\left(\sqrt{\varpi/R_d}\right)}{T^{1/2}}\left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{500\,\mathrm{pc}}{R_d}\right)^{1/2}\left(\frac{100\,\mathrm{pc}}{\varpi}\right)^{5/2}\ \mathrm{W\,m^{-3}} \\ \Gamma_{1,B1,\text{molecular}} &\sim \frac{3\cdot 39\times 10^{-23}\, a\, \cos^{-1}\left(\sqrt{\varpi/R_d}\right)}{T^{1/2}}\left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{500\,\mathrm{pc}}{R_d}\right)^{1/2}\left(\frac{100\,\mathrm{pc}}{\varpi}\right)^{5/2}\ \mathrm{W\,m^{-3}}\end{aligned}\right\} \quad (53\text{-}B1)$$

$$\left.\begin{aligned}\Gamma_{1,B2,\text{atomic}} &\sim \frac{6\cdot 32\times 10^{-23}\, a\left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^2}{T}\left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{M_d}{10^{7} M_\odot}\right)\left(\frac{500\,\mathrm{pc}}{R_d}\right)\left(\frac{100\,\mathrm{pc}}{\varpi}\right)^{3}\ \mathrm{W\,m^{-3}} \\ \Gamma_{1,B2,\text{molecular}} &\sim \frac{1\cdot 22\times 10^{-22}\, a\left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^2}{T}\left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{M_d}{10^{7} M_\odot}\right)\left(\frac{500\,\mathrm{pc}}{R_d}\right)\left(\frac{100\,\mathrm{pc}}{\varpi}\right)^{3}\ \mathrm{W\,m^{-3}}\end{aligned}\right\} \quad (53\text{-}B2)$$

(The heating rate per unit volume for a molecular disc is slightly higher than that for an atomic disc at the same temperature because the molecular disc has a slightly smaller volume, due to the dependence of v_s on $\bar{m}$).

We now compare this heat source (infall) with the other three that have been discussed. First consider Γ_2 (defined by equation (48)). Assuming $M_n = 10^{10}\ M_\odot$, $R_n = 1$ kpc and $\bar{m}_*/\bar{L}_* \sim 3$ (cf. Sanders & Lowinger, 1972), this becomes

$$\Gamma_2 \sim 3\cdot 5\times 10^{-22}\, f_2 \left(\frac{100\,\mathrm{pc}}{\varpi}\right)^2 \quad \mathrm{W\,m^{-3}} \qquad (54)$$

At first sight it might appear that $\Gamma_2 \gtrsim \Gamma_1$, but if this were the case the factor f_2 would have to be of order unity. Even in

H II regions the efficiency of conversion of stellar radiation into kinetic energy is not high (being typically $\lesssim 1\ \%$; Bohuski, 1973), and under the circumstances of the present problem (where the disc is embedded in a cluster of predominantly population II stars) it is probably substantially less than this value. This conclusion may be strengthened by noting that starlight is absorbed mainly by dust, and in order for energy to be transferred from the dust to the gas it is necessary that $T \lesssim T_{dust}$. Under normal interstellar conditions dust temperatures are very low (typically ≈ 20 K; Leung, 1975), so starlight could only be an important heat source if the gas was extremely cold. However interstellar gases at such low temperatures radiate energy extremely inefficiently, and in view of the strong mechanical heating due to infalling gas it is questionable whether the temperature could ever become low enough for dust heating to be important. It is shown later, in fact, that in all circumstances where dust-gas collisions are of importance to the energetics of the gas, $T_{dust} \lesssim T$, so in the present problem dust always acts as a coolant rather than as a heat source. It is therefore concluded that $\Gamma_1 \gg \Gamma_2$.

The heating effect due to supernovae within the disc is given by (49). Adopting $M_n \sim 10^{10}\ M_\odot$, $\tau \sim 10^4$ years per $10^{10}\ M_\odot$ (Tammann, 1974) and $\epsilon_{SN} \sim 10^{44}$ J (e.g. Mathews & Baker, 1971), this becomes

$$\Gamma_3 \sim 8.6 \times 10^{-26} \left(\frac{100\,pc}{\varpi}\right)^2 \quad W\,m^{-3} \qquad (55)$$

Thus, provided $T \lesssim 10^4$ K, $\Gamma_3 \lesssim \Gamma_1$

Finally, consider the ratio Γ_4 / Γ_1 . This may be obtained by comparing (51) with (47) and gives

$$\Gamma_4/\Gamma_1 \sim \frac{10\, G\, \bar{m}_* \, M_d}{a\, \varpi^2\, V_0\, \alpha_*\, M_n}$$

Adopting $\bar{m}_* \sim 1\ M_\odot$, $\alpha_* M_n \sim 10^{-2}\ M_\odot\ y^{-1}$ and $V_0 \sim 207\ \mathrm{km\ s^{-1}}$, we obtain

$$\Gamma_4/\Gamma_1 \sim \frac{2\times 10^{-5}}{a} \left(\frac{100\,\mathrm{pc}}{\varpi}\right)^2 \left(\frac{M_d}{10^7 M_\odot}\right) \qquad (56)$$

so it is concluded that heating by the passage of stars through the disc is also negligible compared to Γ_1 , except possibly, in the case of massive discs, very close to the centre.

The most important energy source for the disc is therefore heating by infalling material, and the total volumetric heating rate may be estimated from (53). For convenience this is re-written in the following form:

$$\left.\begin{aligned} \Gamma_{A1,\,\mathrm{atomic}} &\sim \frac{2\times 10^{-23} \cos^{-1}(\varpi/R_d)}{T^{1/2}} \left(\frac{100\,\mathrm{pc}}{\varpi}\right)^2 \gamma_{A1} \quad \mathrm{W\,m^{-3}} \\ \Gamma_{A1,\,\mathrm{molecular}} &\sim \frac{3\times 10^{-23} \cos^{-1}(\varpi/R_d)}{T^{1/2}} \left(\frac{100\,\mathrm{pc}}{\varpi}\right)^2 \gamma_{A1} \quad \mathrm{W\,m^{-3}} \end{aligned}\right\} \quad (57\text{-}A1)$$

$$\left.\begin{aligned} \Gamma_{A2,\,\mathrm{atomic}} &\sim \frac{5\times 10^{-23} \left(\cos^{-1}(\varpi/R_d)\right)^2}{T} \left(\frac{100\,\mathrm{pc}}{\varpi}\right)^2 \gamma_{A2} \quad \mathrm{W\,m^{-3}} \\ \Gamma_{A2,\,\mathrm{molecular}} &\sim \frac{10^{-22} \left(\cos^{-1}(\varpi/R_d)\right)^2}{T} \left(\frac{100\,\mathrm{pc}}{\varpi}\right)^2 \gamma_{A2} \quad \mathrm{W\,m^{-3}} \end{aligned}\right\} \quad (57\text{-}A2)$$

$$\left.\begin{aligned} \Gamma_{B1,\,atomic} &\sim \frac{2\times 10^{-23}\cos^{-1}\left(\sqrt{\varpi/R_d}\right)}{T^{1/2}}\left(\frac{100\,pc}{\varpi}\right)^{5/2}\gamma_{B1} \quad W\,m^{-3} \\ \Gamma_{B1,\,molecular} &\sim \frac{3\times 10^{-23}\cos^{-1}\left(\sqrt{\varpi/R_d}\right)}{T^{1/2}}\left(\frac{100\,pc}{\varpi}\right)^{5/2}\gamma_{B1} \quad W\,m^{-3} \end{aligned}\right\} \quad (57\text{-}B1)$$

$$\left.\begin{aligned} \Gamma_{B2,\,atomic} &\sim \frac{6\times 10^{-23}\left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^2}{T}\left(\frac{100\,pc}{\varpi}\right)^{3}\gamma_{B2} \quad W\,m^{-3} \\ \Gamma_{B2,\,molecular} &\sim \frac{10^{-22}\left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^2}{T}\left(\frac{100\,pc}{\varpi}\right)^{3}\gamma_{B2} \quad W\,m^{-3} \end{aligned}\right\} \quad (57\text{-}B2)$$

where the various constants γ depend on particular details of the model and are defined by

$$\gamma_{A1} = a\left(\frac{\alpha_* \mathcal{M}_*}{10^{-2} M_\odot\, y^{-1}}\right)\left(\frac{500\,pc}{R_d}\right) \qquad (58\text{-}A1)$$

$$\gamma_{A2} = a\left(\frac{\alpha_* \mathcal{M}_*}{10^{-2} M_\odot\, y^{-1}}\right)\left(\frac{500\,pc}{R_d}\right)^2\left(\frac{\mathcal{M}_d}{10^{7} M_\odot}\right) \qquad (58\text{-}A2)$$

$$\gamma_{B1} = a\left(\frac{\alpha_* \mathcal{M}_*}{10^{-2} M_\odot\, y^{-1}}\right)\left(\frac{500\,pc}{R_d}\right)^{1/2} \qquad (58\text{-}B1)$$

$$\gamma_{B2} = a\left(\frac{\alpha_* \mathcal{M}_*}{10^{-2} M_\odot\, y^{-1}}\right)\left(\frac{500\,pc}{R_d}\right)\left(\frac{\mathcal{M}_d}{10^{7} M_\odot}\right) \qquad (58\text{-}B2)$$

V-6 Cooling processes

The cooling rate per unit volume may be written under most circumstances in the form

$$L = n^2 \Lambda(T) \qquad (59)$$

where n is the particle number density and Λ the usual cooling function. Graphs of Λ appropriate for an atomic gas of normal cosmic abundances and specified fractional ionisation have been published by a number of authors (e.g. Dalgarno & McCray, 1972; Schwarz et al, 1972; Falle, 1975), and values of Λ ($W\ m^3$) estimated from figure 1 of Schwarz et al are tabulated below for two assumed values of the fractional ionisation x.

T (K)	$\Lambda(x=10^{-3}, T)$	$\Lambda(x=10^{-1}, T)$
10	$\sim 10^{-43}$	$\sim 10^{-41}$
20	$\sim 10^{-41}$	6×10^{-40}
30	4×10^{-41}	2×10^{-39}
40	8×10^{-41}	$3{\cdot}5 \times 10^{-39}$
50	$1{\cdot}2 \times 10^{-40}$	5×10^{-39}
100	4×10^{-40}	10^{-38}
200	9×10^{-40}	2×10^{-38}
300	$1{\cdot}3 \times 10^{-39}$	$2{\cdot}5 \times 10^{-38}$
400	2×10^{-39}	3×10^{-38}
1000	$3{\cdot}5 \times 10^{-39}$	4×10^{-38}
3000	5×10^{-39}	4×10^{-38}
5000	6×10^{-39}	3×10^{-38}
10000	8×10^{-39}	10^{-37}

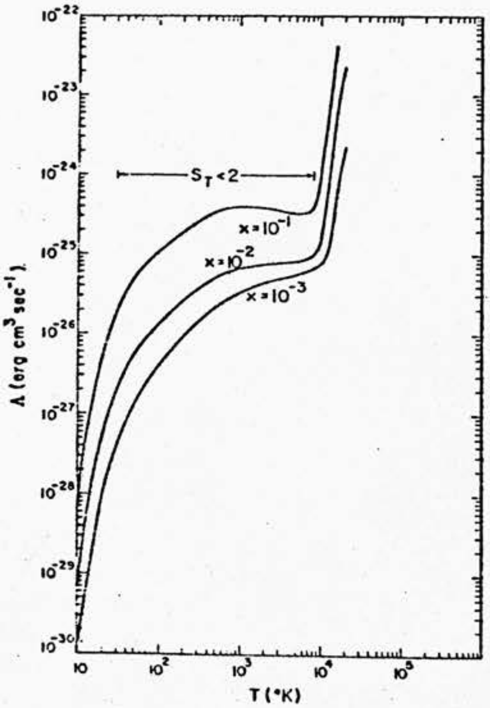


FIG. 1.—Cooling of a low-density plasma of cosmic abundance. The power radiated per unit mass is given by $P = n_0\Lambda/m_H$. For $T < 10^4$ °K, Λ is given for three values of $x = n_e/n_0$. Temperature ranges favoring thermal condensation (logarithmic slope $s_T < 2$) are indicated.

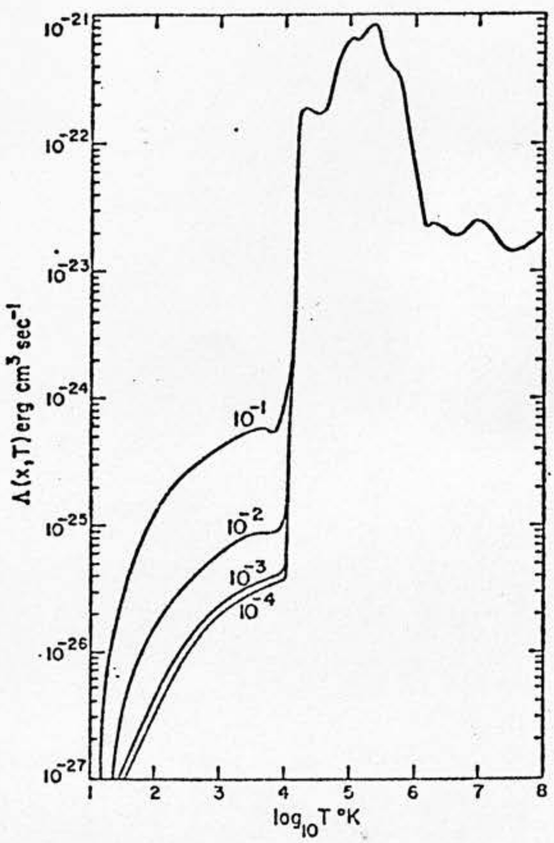


FIGURE 2. The interstellar cooling function $\Lambda(x, T)$ for various values of the fractional ionization x. The labels refer to the values of x.

For convenience, graphs of the cooling function have been reproduced above, the upper figure taken from Schwarz et al (1972) and the lower one from Dalgarno & McCray (1972). For $x \lesssim 10^{-3}$, Λ is almost independent of the fractional ionisation, because hydrogen impact excitations control the cooling, but for $x \gtrsim 10^{-2}$, Λ increases almost linearly with x. Inspection of Dalgarno & McCray's cooling curve shows that at $T \sim 10^4$ K, Λ increases very rapidly with temperature, and reaches a local maximum of order 2×10^{-35} W m^3 at $T \sim 2 \times 10^4$ K. The sharp increase in cooling efficiency at this temperature means that the disc temperature may be expected to be $\lesssim 10^4$ K. This expectation is confirmed quantitatively in Section V-8.

At low temperatures and normal densities cooling of interstellar gases is dominated by fine structure transitions of singly-ionised carbon (C II). If for any reason this should be depleted (for example by accretion of carbon onto grains or by ionisation to C III) the cooling function will be modified and may be significantly less than the values tabulated here (cf. Dalgarno & McCray, 1972; Figure 3). On the other hand, in dense regions the radiation which would otherwise ionise carbon might be greatly attenuated, thereby allowing a significant fractional abundance of neutral carbon to form. Cooling by fine structure transitions of C I at low temperatures is much more efficient than cooling by C II (by factors ranging from about 5 at $T \sim 20$ K, to about 200 at $T \sim 10$ K; Dalgarno & McCray, 1972, Table 4), and consequently the presence of even a small fraction of neutral carbon might lead to a significant <u>increase</u> in Λ at low temperatures. An added complication here is that in cold dense regions the composition of the gas might be altered by the formation of molecules. For these reasons a realistic determination of the appropriate cooling function at low temperatures is an extremely difficult problem, and it is possible that the temperature estimates that are made below might be subject to a systematic error.

If the gas does condense into molecular form, cooling at moderate densities is dominated by radiation from the CO molecule. This cooling has been calculated by Scoville & Solomon (1974), who obtained the following analytic approximation to the CO cooling rate per unit volume:

$$L_{CO} \sim \begin{cases} 6\times 10^{-42}\, n_{CO}\, n_{H_2}\, T^2 \quad W\, m^{-3} & n_{CO} n_{H_2} \lesssim 10^{14}\, m^{-6} \\ 2\times 10^{-28}\, T^3 \quad W\, m^{-3} & n_{CO} n_{H_2} \gg 10^{14}\, m^{-6} \end{cases} \qquad (60)$$

If n_{CO}/n_{H_2} is assumed to be $\sim 10^{-4}$, this becomes

$$L_{CO} \sim \begin{cases} 6\times 10^{-46}\, T^2\, n_{H_2}^2 \quad W\, m^{-3} & n_{H_2} \leq 10^9\, m^{-3} \\ 2\times 10^{-28}\, T^3 \quad W\, m^{-3} & n_{H_2} \gg 10^9\, m^{-3} \end{cases} \qquad (61)$$

Recent work by Goldsmith & Langer (1978) has to a certain extent superseded the work by Scoville & Solomon referred to above, and although cooling at densities $\lesssim 10^9\ m^{-3}$ is still dominated by the ^{12}CO molecule, at higher densities ($10^9 \lesssim n_{H_2} \lesssim 10^{11}\ m^{-3}$) rarer isotopes of CO, and C I and O_2 are found together to contribute between 30 and 70 % towards the total cooling. At very high densities ($n_{H_2} \gtrsim$ few x $10^{11}\ m^{-3}$) cooling is dominated by H_2O and in addition other coolants are found to be comparable in strength to CO. Although the volumetric heating rate used here (equation (61)) thus probably underestimates the high-density cooling by a factor possibly as large as 10, inclusion of these more complicated effects only strengthens the general conclusion (Section V-8) that close to the centre the disc is very cold.

At high densities inelastic collisions between gas molecules and dust grains become an important coolant (for $T \gtrsim T_d$), and it may be shown (e.g. Spitzer, 1949; Leung, 1975) that the cooling rate due to this process is

$$L_{g-d} \sim \frac{3}{2} k \Delta T \, \xi \left(\frac{8kT}{\pi \bar{m}}\right)^{1/2} n_{H_2} n_d \, \pi \bar{a}^2 \qquad (62)$$

Here $\Delta T = T - T_d$ is the difference between gas and dust temperatures, ξ is a factor of order unity describing the degree of elasticity of the collisions, $\bar{a}$ is the characteristic grain size, and $\bar{m}$ is the mean particle mass. $\bar{m}$ is taken to be $2{\cdot}5\ m_H$ (appropriate for a molecular system), and using the same other parameter values as Leung (1975): $\xi = 0{\cdot}5$, $\bar{a} = 0{\cdot}06\ \mu m$, $n_d/n_{H_2} = 2 \times 10^{-11}$, (62) becomes

$$L_{g-d} \sim 2{\cdot}1 \times 10^{-46}\ n_{H_2}^2\ T^{1/2} (T - T_d) \quad W\,m^{-3} \qquad (63)$$

Further progress requires knowledge of the dust temperature, which is determined by the balance between the energy absorbed by a grain and that radiated at far-infrared wavelengths at temperature T_d. An exact solution to this difficult radiative transfer problem is beyond the scope of the present investigation (see for example Werner & Salpeter, 1969; Greenberg, 1971; Leung, 1975); and T_d is estimated in the next section by approximate methods.

V-7 Estimated dust temperature

Dust grains in radiation fields typical of the solar neighbourhood have temperatures lying in the broad range $10 \lesssim T_d \lesssim 20$ K (see the references given at the end of Section V-6), depending on the details of the model calculated. In particular Leung (1975) showed that the free-space temperature reached by a typical grain mixture model was $T_d \sim 13$ K, and his Figure 6 showing this result is reproduced below.

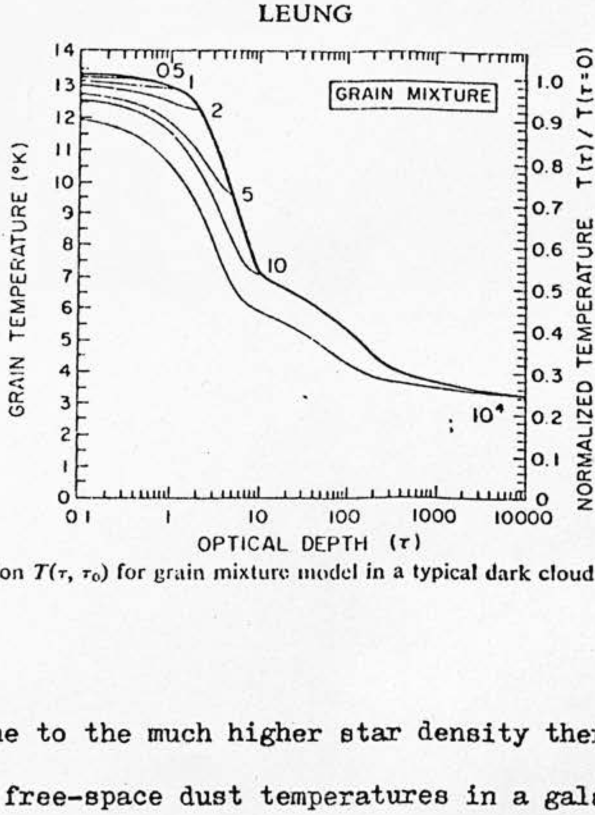


FIG. 6.—Temperature distribution $T(\tau, \tau_0)$ for grain mixture model in a typical dark cloud with a centrally condensed density distribution.

In general, due to the much higher star density there, it might be expected that free-space dust temperatures in a galactic nucleus would be somewhat higher than those in the solar neighbourhood, but because of the strong temperature-dependence of the power radiated by grains in the far-infrared, typically proportional to T_d^{β} where $\beta \sim 6 \pm 1$ (Werner & Salpeter, 1969), a large change in the absorbed energy is required to effect a small change in T_d.

In the present model it is assumed that the ambient radiation field can be approximated by a dilute blackbody distribution with temperature $T_s \sim 4000$ K (Sanders & Lowinger, 1972) and dilution factor η_s. The energy density of such a field is

$$\mathcal{E}_s = \frac{4\sigma}{c} \eta_s T_s^4 \qquad (64)$$

where σ is Stefan's constant, and it may be shown by direct integration of the assumed star distribution that this is also

$$\mathcal{E}_*(\varpi) = \frac{\mathfrak{M}_n}{4\pi R_n} \frac{\bar{L}_*}{\bar{m}_*} \frac{1}{c\varpi} f(\varpi) \qquad (65)$$

where $f(\varpi)$ is a slowly varying function of order 2 defined by

$$f(\varpi) = \frac{\pi^2}{4} - \sum_{k=1}^{\infty} \left(\frac{\varpi}{R_n}\right)^{2k-1} \frac{1}{(2k-1)^2}$$

Equating (64) with (65) and adpoting $f \sim 2$ thus gives

$$\eta_s T_s^4 \sim \frac{\mathfrak{M}_n}{\bar{m}_*} \frac{\bar{L}_*}{8\pi R_n \varpi \sigma}$$

which, with $M_n \sim 10^{10}\ M_\odot$, $R_n \sim 1$ kpc and $\bar{m}_*/\bar{L}_* \sim 3\ M_\odot/L_\odot$ (Sanders & Lowinger, 1972), reduces to

$$\eta_s T_s^4 \sim 10^4 \left(\frac{100\,\mathrm{pc}}{\varpi}\right) \qquad (66)$$

Interstellar grains absorb energy at visual wavelengths with an efficiency $\propto \lambda^{-1}$, which implies that the energy they absorb in the assumed blackbody radiation field is proportional to $\eta_s T_s^5$ (Andriesse, 1974). If the power radiated by grains is assumed to be typically proportional to T_d^6 (cf. the range of β mentioned above), equating absorbed and emitted powers therefore implies

$$T_d^{\,6} \propto \eta_s \, T_s^{\,5}$$

$$\Rightarrow \quad \frac{T_d}{T_{d,o}} \sim \left(\frac{\eta_s \, T_s^{\,4}}{\eta_{s,o} \, T_{s,o}^{\,4}} \cdot \frac{T_s}{T_{s,o}} \right)^{1/6} \qquad (67)$$

where the subscripts zero refer to quantities evaluated under standard solar neighbourhood conditions. The solar neighbourhood radiation field is closely approximated by a dilute blackbody distribution with $T_{s,o} \sim 10^4$ K and $\eta_{s,o} \sim 10^{-14}$ (e.g. Greenberg, 1971), and using these values and also $T_s \sim 4000$ K and $T_{d,o} \sim 13$ K, we obtain the result

$$T_d \sim 24 \left(\frac{100\,\mathrm{pc}}{\varpi} \right)^{1/6} \quad K \qquad (68)$$

This calculation has assumed that the dust is in an unshielded part of the nucleus. In a region of high optical depth dust temperatures will be lower, but it is not possible to obtain precise results without solving the detailed radiative transfer problem that is involved. However inspection of Figure 6 of Leung (1975), reproduced above, shows that the dust temperature at a visual optical depth of order 10 is about half its free-space value, and for $1 \lesssim \tau_v \lesssim 10$ T_d varies approximately linearly with optical depth. At very high optical depths ($\tau_v \gtrsim 200$) Leung found that grain temperatures tended towards a limiting value which depended on the grain model and the radiation field; for the particular grain mixture model referred to above this was found to be close to the temperature of the blackbody distribution with the same energy density as the assumed radiation field. Combining these results, the variation of dust temperature with visual optical depth is estimated in our model to be

$$T_d(\tau_v) \sim \begin{cases} T_d(\tau_v = 0) \cdot \left(1 - \frac{\tau_v}{20}\right) & 0 \lesssim \tau_v \lesssim 10 \\ T_{d,lim} & \tau_v \gtrsim 200 \end{cases} \qquad (69)$$

where the limiting temperature is defined by

$$T_{d,lim} \sim 10 \left(\frac{100\,pc}{\varpi}\right)^{1/4} \; K \qquad (70)$$

Since $T_d(\tau_v \sim 10)$ is very close to the limiting temperature, we have to order of magnitude the result

$$T_d(\tau_v, \varpi) \approx \begin{cases} 25\left(1 - \frac{\tau_v}{20}\right)\left(\frac{100\,pc}{\varpi}\right)^{1/6} \; K & 0 \lesssim \tau_v \lesssim 10 \\ 10 \left(\frac{100\,pc}{\varpi}\right)^{1/4} \; K & \tau_v \gtrsim 10 \end{cases} \qquad (71)$$

The optical depth of a region of length l and number density $\bar{n}_d$ of dust particles may be written, assuming that scattering is isotropic, in the form

$$\tau_v \sim \bar{n}_d \, l \, \langle Q_{ext} \pi a^2 \rangle_v \qquad (72)$$

where $\langle Q_{ext} \pi a^2 \rangle_v$ is the mean extinction cross-section at visual wavelengths. If this is taken to be $\sim 5 \times 10^{-14}$ m^2 (cf. Leung, 1975; Table 3) and the dust/gas ratio is assumed to be $n_d/n_H \sim 10^{-11}$, the visual optical depth to mid-plane becomes

$$\tau_v(\varpi, z=0) \sim \frac{n_d}{n_H} \frac{\sigma(\varpi)}{2\bar{m}} \langle Q_{ext} \pi a^2 \rangle_v$$

$$\Longrightarrow \tau_{V,A}(\varpi,0) \sim 10 \cos^{-1}\left(\frac{\varpi}{R_d}\right)\left(\frac{M_d}{10^7 M_\odot}\right)\left(\frac{500\,pc}{R_d}\right)\left(\frac{100\,pc}{\varpi}\right) \qquad (73-A)$$

$$\tau_{V,B}(\varpi,0) \sim 10 \cos^{-1}\left(\sqrt{\varpi/R_d}\right)\left(\frac{M_d}{10^7 M_\odot}\right)\left(\frac{500\,pc}{R_d}\right)^{1/2}\left(\frac{100\,pc}{\varpi}\right)^{3/2} \qquad (73-B)$$

where A and B as usual refer to Cases A and B respectively. Representative discs with $M_d \gtrsim 10^7\ M_\odot$ thus have visual optical depths $\tau_V \gtrsim 10$ at radii $\varpi \lesssim 100$ pc, and the grain temperatures in the central regions of such systems will be given approximately by

$$T_d(\varpi) \approx 10\left(\frac{100\,pc}{\varpi}\right)^{1/4}\ K \qquad (74)$$

The dust temperature therefore increases inwards close to the centre of the disc. In regions of low optical depth (near the edges of the disc, say) the dust temperature also increases inwards ($T_d \propto \varpi^{-1/6}$, equation (68)), but at intermediate radii where the effects of increasing optical depth dominate over those of increasing star density, T_d will decrease inwards.

V-8 Temperature distribution of the disc

V-8A Proof that infall does not destroy the disc

The volumetric heating rate was calculated in Section V-5 above (equations (57) and (58)); dividing these expressions by n^2 gives an estimate of the cooling function required to just offset this heat source. i.e.

$$\Lambda_{req,A1} \sim \frac{F_{A1}\, a\, T^{1/2}}{\cos^{-1}(\varpi/R_d)} \left(\frac{\alpha_* \mathcal{M}_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500pc}\right)\left(\frac{\varpi}{100pc}\right)^2\left(\frac{10^7 M_\odot}{\mathcal{M}_d}\right)^2 \quad W m^3 \qquad (75\text{-}A1)$$

$$\Lambda_{req,A2} \sim \frac{F_{A2}\, a\, T}{\left(\cos^{-1}(\varpi/R_d)\right)^2} \left(\frac{\alpha_* \mathcal{M}_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500pc}\right)^2\left(\frac{\varpi}{100pc}\right)^2\left(\frac{10^7 M_\odot}{\mathcal{M}_d}\right)^3 \quad W m^3 \qquad (75\text{-}A2)$$

$$\Lambda_{req,B1} \sim \frac{F_{B1}\, a\, T^{1/2}}{\cos^{-1}(\sqrt{\varpi/R_d})} \left(\frac{\alpha_* \mathcal{M}_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500pc}\right)^{1/2}\left(\frac{\varpi}{100pc}\right)^{5/2}\left(\frac{10^7 M_\odot}{\mathcal{M}_d}\right)^2 \quad W m^3 \qquad (75\text{-}B1)$$

$$\Lambda_{req,B2} \sim \frac{F_{B2}\, a\, T}{\left(\cos^{-1}(\sqrt{\varpi/R_d})\right)^2} \left(\frac{\alpha_* \mathcal{M}_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500pc}\right)\left(\frac{\varpi}{100pc}\right)^3\left(\frac{10^7 M_\odot}{\mathcal{M}_d}\right)^3 \quad W m^3 \qquad (75\text{-}B2)$$

where the constants F are defined by

$$F_{A1} \sim \begin{cases} 4 \times 10^{-43} & \text{for an atomic disc} \\ 10^{-42} & \text{for a molecular disc} \end{cases}$$

$$F_{A2} \sim \begin{cases} 2 \times 10^{-43} & \text{for an atomic disc} \\ 4 \times 10^{-43} & \text{for a molecular disc} \end{cases}$$

$$F_{B1} \sim \begin{cases} 3 \times 10^{-43} & \text{for an atomic disc} \\ 9 \times 10^{-43} & \text{for a molecular disc} \end{cases}$$

$$F_{B2} \sim \begin{cases} 1 \times 10^{-43} & \text{for an atomic disc} \\ 2 \times 10^{-43} & \text{for a molecular disc} \end{cases}$$

The disc temperature may be obtained from these results for any particular set of the parameters by finding the value of T at which Λ_{req} equals the value of $\Lambda(x,T)$ listed earlier (p.79). Before tackling this problem however, we first demonstrate that the disc can absorb the infalling gas without being destroyed.

The condition that infalling gas does not heat the disc to unacceptably high temperatures is essentially the condition that the equilibrium temperature obtained by equating Λ with Λ_{req} falls to the left of the maximum in the cooling function which occurs in the temperature range $10^5 - 10^6$ K (see the cooling curve graphed by Dalgarno & McCray (1972), reproduced here on page 80). If Λ_{req} exceeds this peak value (which is on the order of $\sim 5 \times 10^{-35}$ W m^3) no low-temperature equilibrium will be possible, and the infalling gas would cause the disc to heat up, eventually producing a hot outflowing galactic wind. A rough estimate of the region of parameter space within which infalling material does not destroy the disc is obtained by setting

$$\Lambda_{req}\left(T \sim 10^5 K\right) \lesssim \Lambda_{max} \sim 5 \times 10^{-35} \text{ W m}^3 \qquad (76)$$

It is thus concluded that the disc does act as a sink for the infalling material provided:

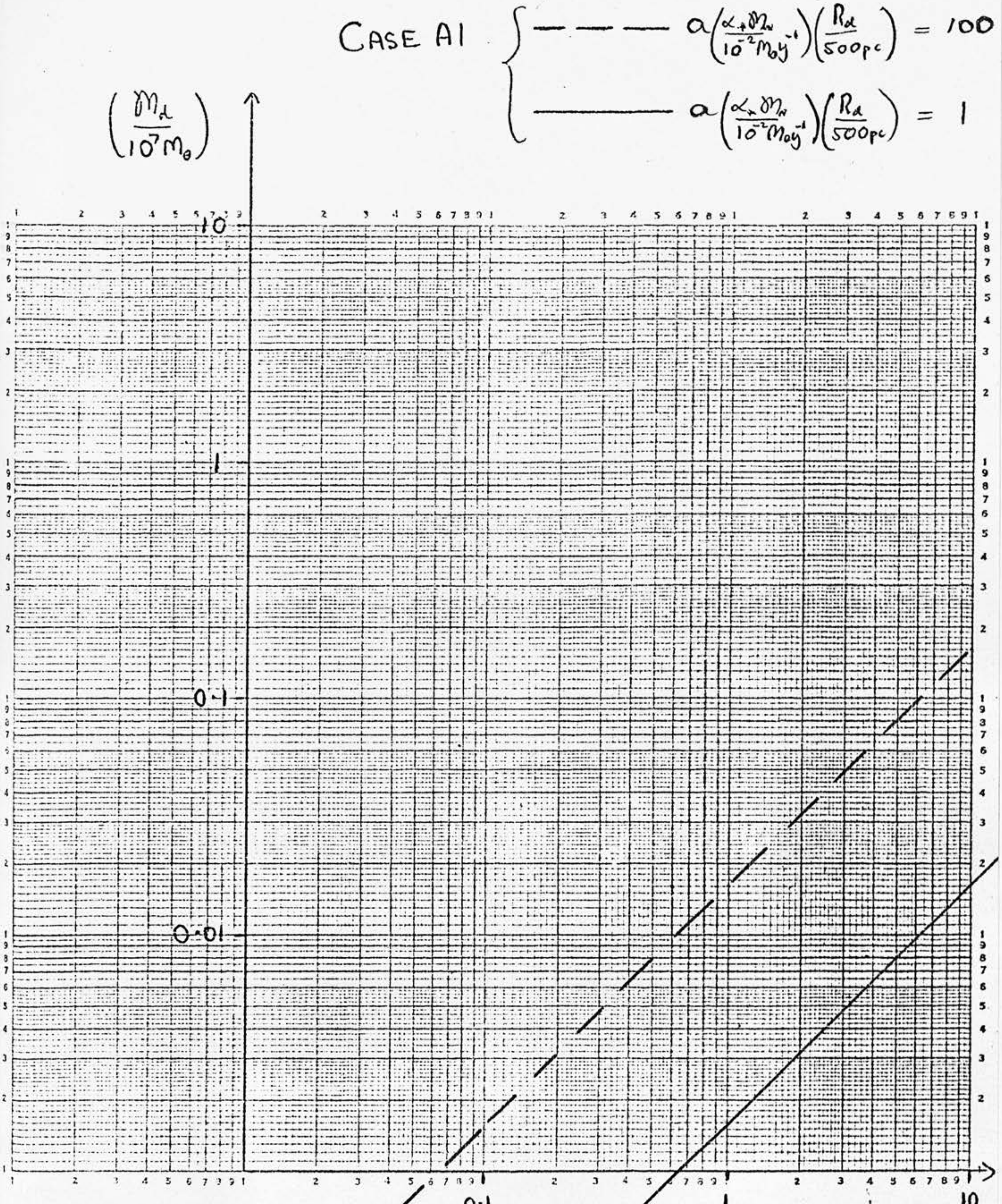
$$\left(\frac{M_d}{10^7 M_\odot}\right)_{A1} \gtrsim \left[\frac{F_{A1}\, a\, 10^{5/2}}{\cos^{-1}(\varpi/R_d)} \cdot \left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,pc}\right)\frac{1}{5\times10^{-35}}\right]^{1/2}\left(\frac{\varpi}{100\,pc}\right) \qquad (77\text{-}A1)$$

$$\left(\frac{M_d}{10^7 M_\odot}\right)_{A2} \gtrsim \left[\frac{F_{A2}\, a\, 10^{5}}{\left(\cos^{-1}(\varpi/R_d)\right)^2}\left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,pc}\right)^2\frac{1}{5\times10^{-35}}\right]^{1/3}\left(\frac{\varpi}{100\,pc}\right)^{2/3} \qquad (77\text{-}A2)$$

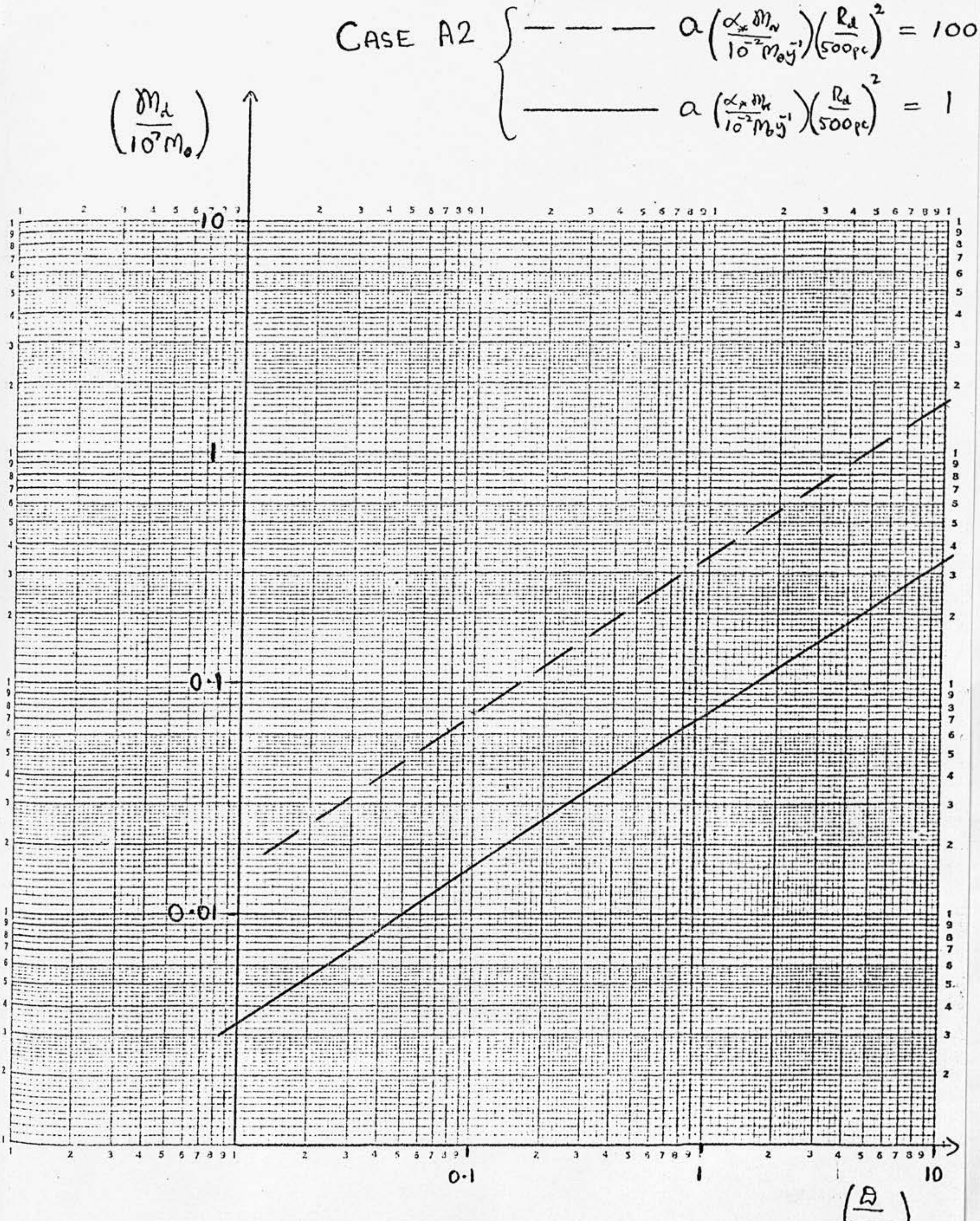
$$\left(\frac{M_d}{10^7 M_\odot}\right)_{B1} \gtrsim \left[\frac{F_{B1}\, a\, 10^{5/2}}{\cos^{-1}(\sqrt{\varpi/R_d})}\left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,pc}\right)^{1/2}\frac{1}{5\times10^{-35}}\right]^{1/2}\left(\frac{\varpi}{100\,pc}\right)^{5/4} \qquad (77\text{-}B1)$$

$$\left(\frac{M_d}{10^7 M_\odot}\right)_{B2} \gtrsim \left[\frac{F_{B2}\, a\, 10^{5}}{\left(\cos^{-1}(\sqrt{\varpi/R_d})\right)^2}\left(\frac{\alpha_* M_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,pc}\right)\frac{1}{5\times10^{-35}}\right]^{1/3}\left(\frac{\varpi}{100\,pc}\right) \qquad (77\text{-}B2)$$

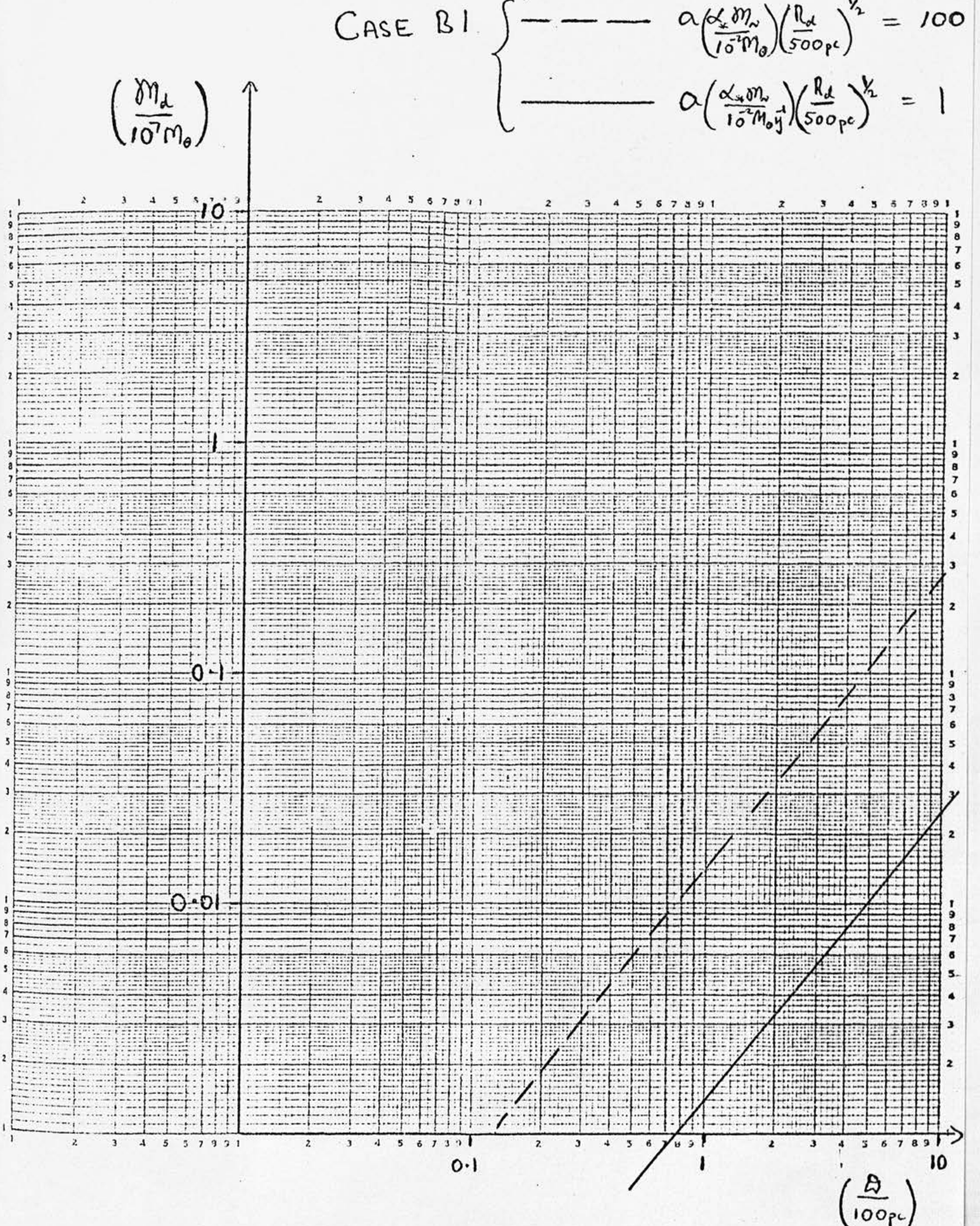
Throughout the inner regions of the disc ($\varpi \lesssim \frac{1}{2} R_d$, say) the inverse cosine parts of these formulae are slowly varying and of order unity. If these terms are set equal to unity, the curves dividing 'allowable' regions of parameter space from non-allowable parts become straight lines in the $(\log M_d - \log \varpi)$-plane, and the four graphs drawn on the next pages show how the criteria (77) depend on disc mass and radius. The region of the $(\log M_d - \log \varpi)$-plane within which discs are not destroyed by infalling material (the 'allowable' region) lies to the left and above the lines drawn, and the boundary has been drawn in each case for two extreme values of the assumed heating function. It should be emphasised that because of the neglect of the inverse cosine variations the actual boundary differs radically from those drawn for ϖ close to R_d, but in spite of this it is possible to conclude that massive discs ($M_d \gtrsim 10^7\ M_\odot$, say) can absorb even hot infalling material without being heated to an unacceptably high temperature.



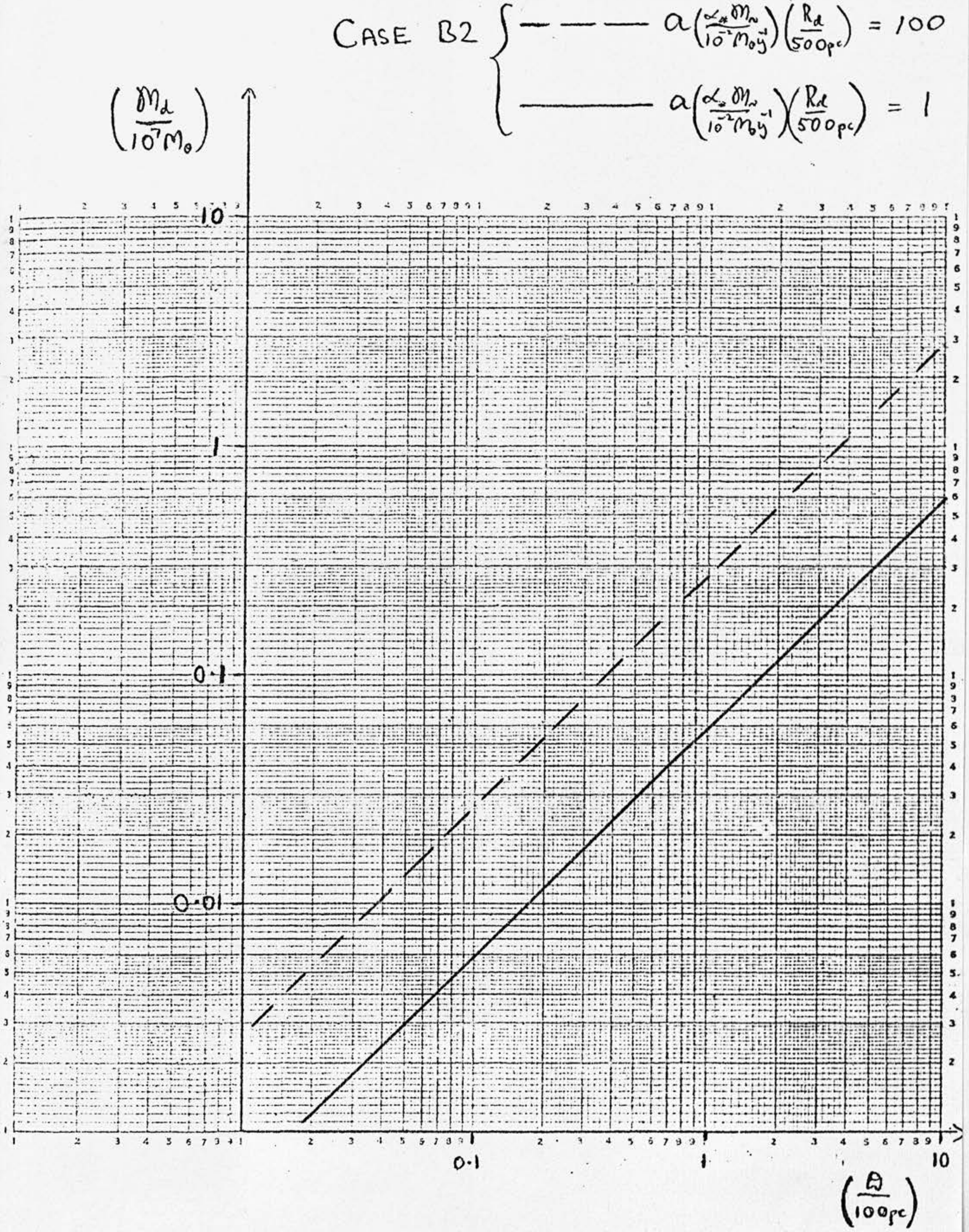
Loci representing the limit of low-temperature equilibrium at $T \sim 10^5$ K



Loci representing the limit of low-temperature equilibrium at $T \sim 10^5$ K



Loci representing the limit of low-temperature equilibrium at $T \sim 10^5$ K



Loci representing the limit of low-temperature equilibrium at $T \sim 10^5$ K

V-8B Temperature of a mainly atomic disc

The equilibrium disc temperature can be obtained, as was noted above, by equating the actual cooling function $\Lambda(x,T)$ to that required to just offset the calculated heating rate. Unfortunately, because Λ is sensitively dependent on the assumed fractional ionisation x (see the discussion in Section V-6), and the radial variation of this quantity has not been established, the precise radial temperature distribution in a particular disc is difficult to obtain accurately. In order to try and bracket the likely range of disc temperatures, loci of constant T have therefore been drawn for the following two cases:

(a) $a\left(\frac{\dot{M}_n}{10^{-2} M_\odot y^{-1}}\right) \sim 10; \quad x \sim 10^{-3}$

(b) $a\left(\frac{\dot{M}_n}{10^{-2} M_\odot y^{-1}}\right) \sim 1; \quad x \sim 10^{-1}$

The first of these corresponds to the case of hot ($T \sim 10^7$ K) inflowing matter falling onto an atomic disc of low fractional ionisation (so cooling is relatively inefficient); and the second corresponds to the case of relatively cold gas being absorbed by a disc with somewhat more efficient cooling. Equating $\Lambda(x,T)$ with Λ_{req} therefore gives:

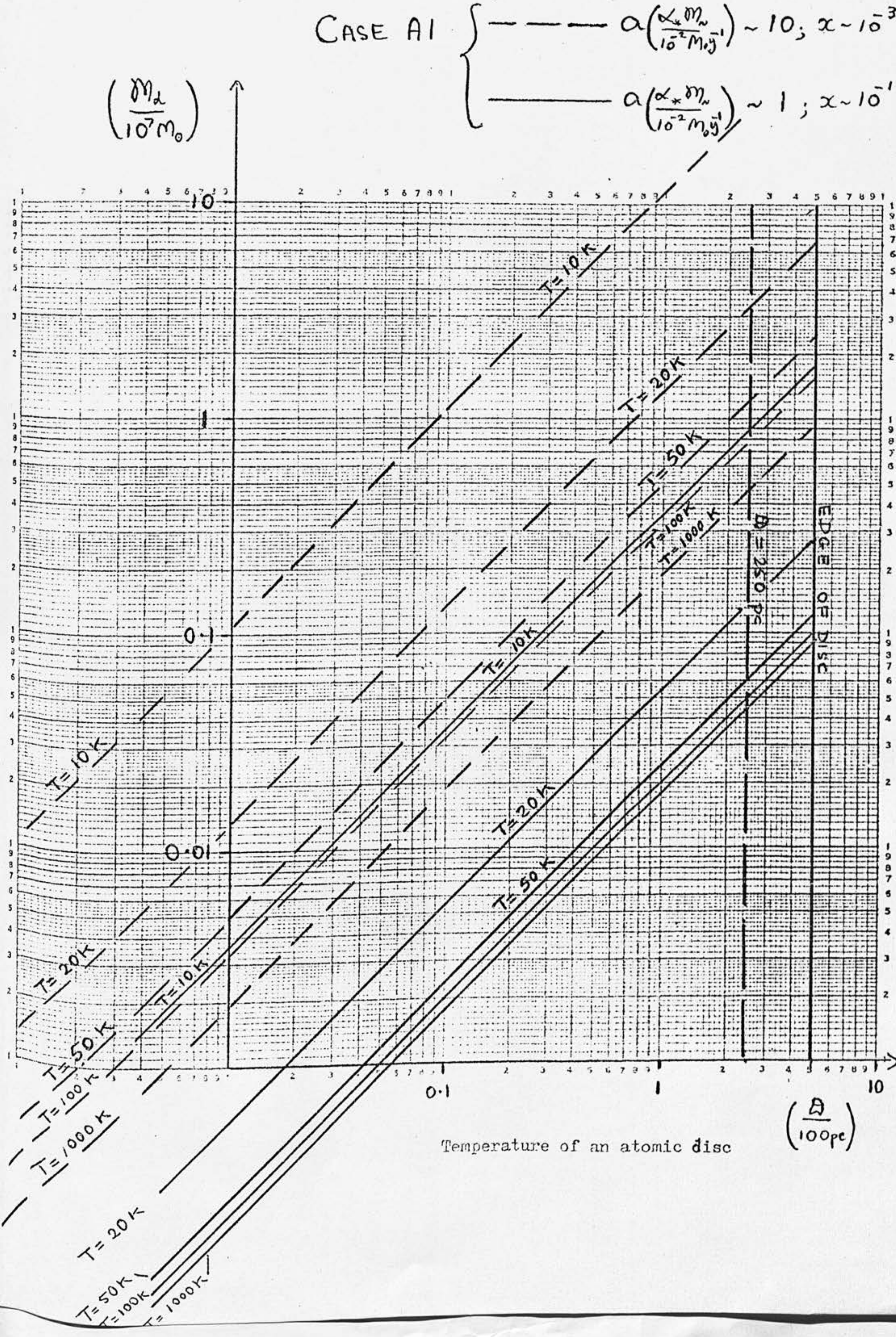
$$\left(\frac{M_d}{10^7 M_\odot}\right)_{A1} \sim \left[\frac{4\times10^{-43}}{\Lambda T^{-1/2}} \cdot a\left(\frac{\alpha_* \dot{M}_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,pc}\right)\right]^{1/2} \cdot \left(\frac{\varpi}{100\,pc}\right)\frac{1}{\left(\cos^{-1}(\varpi/R_d)\right)^{1/2}} \qquad (78\text{-}A1)$$

$$\left(\frac{M_d}{10^7 M_\odot}\right)_{A2} \sim \left[\frac{2\times10^{-43}}{\Lambda T^{-1}} \cdot a\left(\frac{\alpha_* \dot{M}_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,pc}\right)^2\right]^{1/3} \cdot \left(\frac{\varpi}{100\,pc}\right)^{2/3}\frac{1}{\left(\cos^{-1}(\varpi/R_d)\right)^{2/3}} \qquad (78\text{-}A2)$$

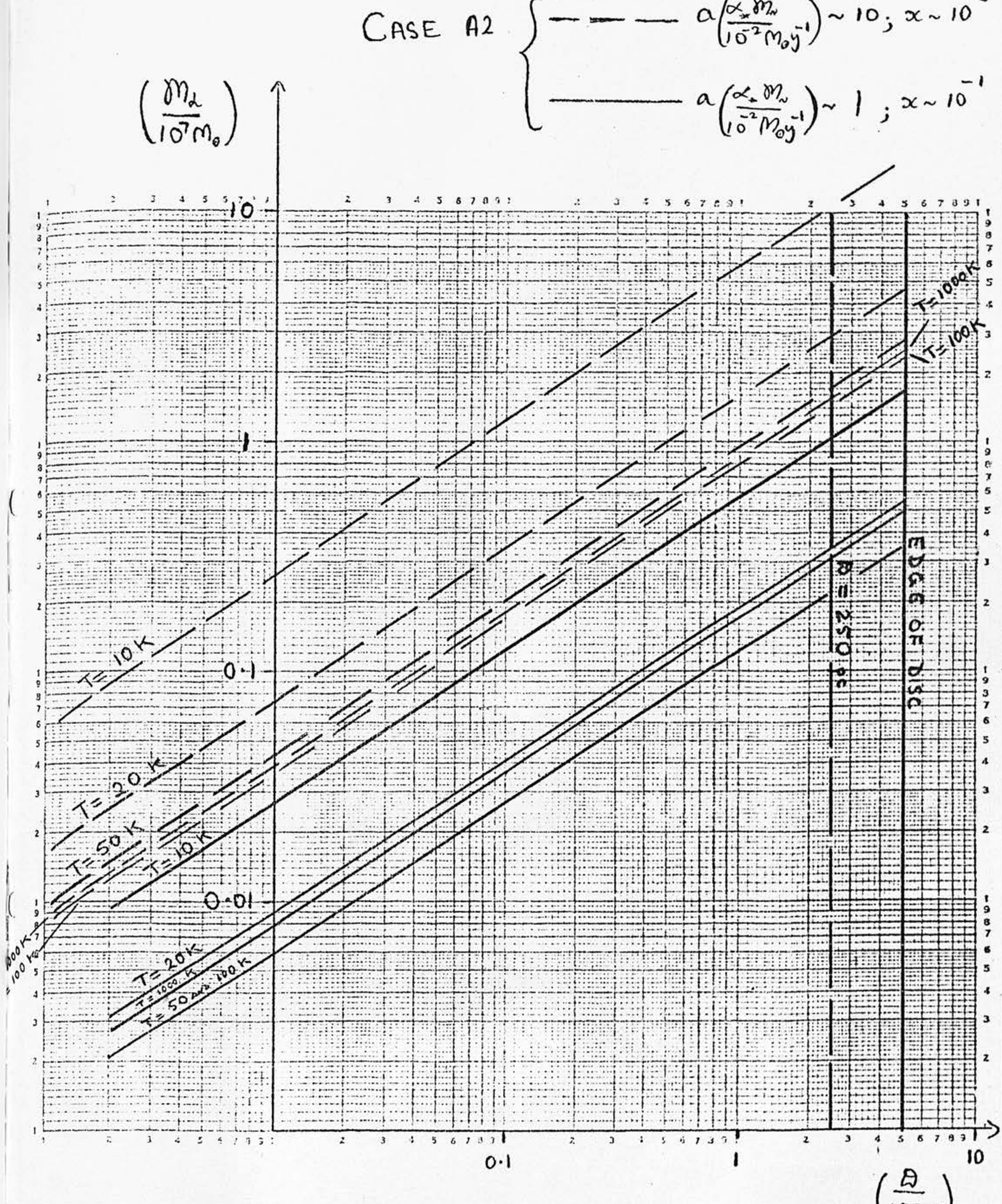
$$\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)_{B1} \sim \left[\frac{3\times10^{-43}}{\Lambda T^{-1/2}} \cdot a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,\text{pc}}\right)^{1/2}\right]^{1/2} \cdot \left(\frac{\varpi}{100\,\text{pc}}\right)^{5/4} \frac{1}{\left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^{1/2}} \qquad (78-B1)$$

$$\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)_{B2} \sim \left[\frac{1\times10^{-43}}{\Lambda T^{-1}} \cdot a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,\text{pc}}\right)\right]^{1/3} \cdot \left(\frac{\varpi}{100\,\text{pc}}\right) \frac{1}{\left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^{2/3}} \qquad (78-B2)$$

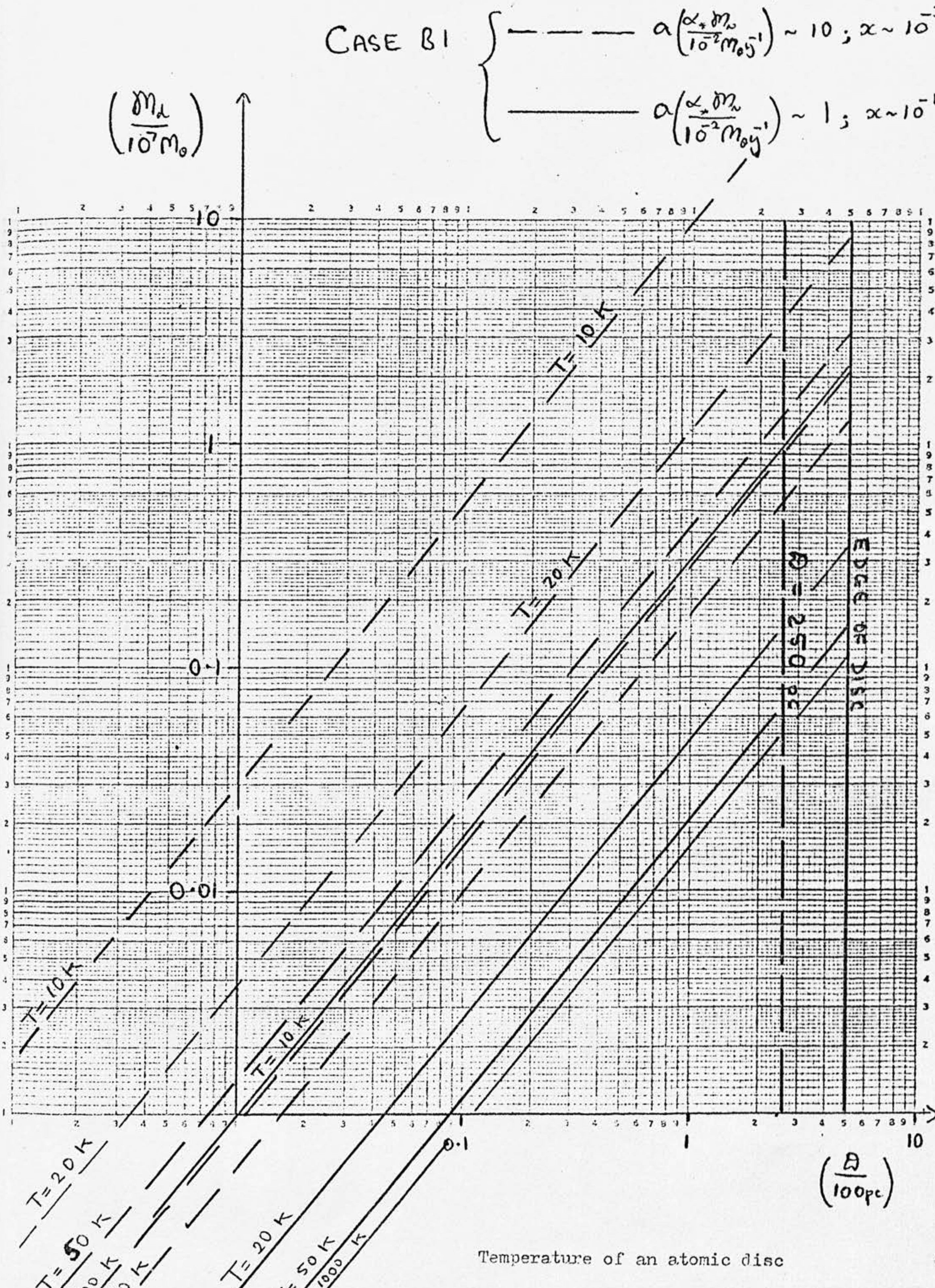
The graphs drawn on the following pages have each been calculated (for clarity) on the assumption that $R_d = 500$ pc, and since their main purpose is just to give an indication as to the general trend of the temperature distribution, the inverse cosine variations have again been assumed to be of order unity. This means that the curves drawn cease to be a good representation of equations (78) for radii $\varpi \gtrsim 250$ pc. (The exact curves tend to rise above those drawn in the outer parts of the disc, $\varpi \gtrsim 250$ pc say, in such a way that $M_d \rightarrow \infty$ as $\varpi \rightarrow R_d = 500$ pc). The values used for $\dot{\Lambda}(x,T)$ were obtained from the table on page 79.



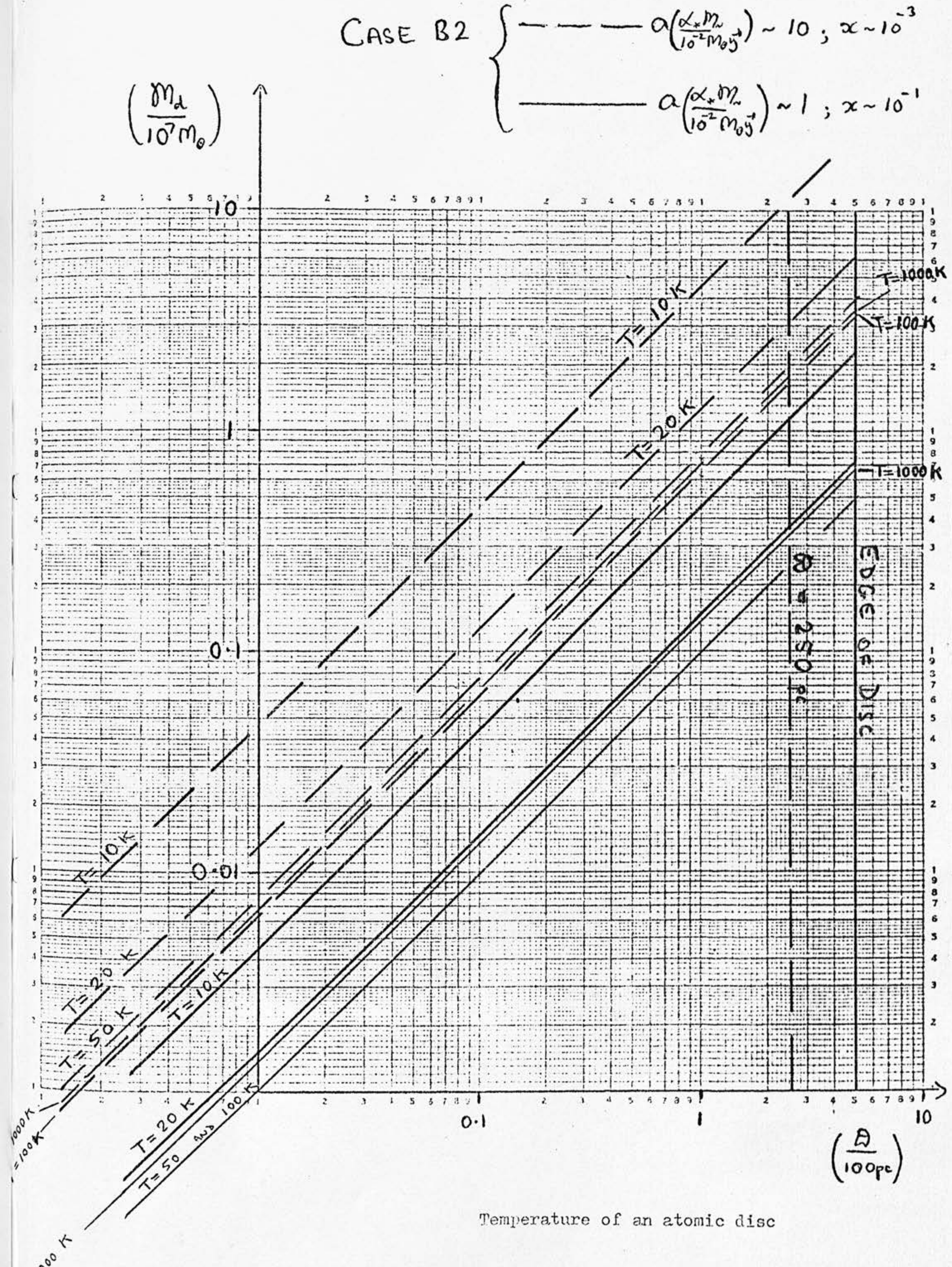
Temperature of an atomic disc



Temperature of an atomic disc



Temperature of an atomic disc



Temperature of an atomic disc

Discussion

The principal conclusion to be drawn from the temperature profiles indicated by these graphs is that massive discs ($M_d \gtrsim 10^7\ M_\odot$) are extremely cold in their central regions, and in the case of self-gravity negligible (Cases A1 and B1) even hot infalling material does not lead to equilibrium temperatures higher than ~ 20 K within the central ~ 100 pc of the disc. Whether or not self-gravity is important at radius ϖ for a given disc must be checked for each individual case, and it is not possible to conclude generally, for example by comparing the graphs for Cases A1 and A2, that self-gravitating discs are on average hotter than non self-gravitating ones. It will be shown later that Case B discs might be significantly less massive than Case A ones (owing to the earlier onset of gravitational instability in these more centrally condensed systems), but even for $M_d \sim 10^5\ M_\odot$ and the case of hot infall the central few parsecs of Case B discs still cool towards equilibrium temperatures $\lesssim 50$ K. In the case of cold infalling material this conclusion applies out to a radius of order 50 pc. It is therefore concluded generally that unless an important heat source has been omitted from the theory, the central regions of an atomic disc will be extremely cold.

In view of this result it is important to investigate the temperature distribution of a molecular disc, but before turning to this question we first verify that the effect of possible dust heating - ignored in the above estimation of T - is indeed negligible even at very low atomic gas temperatures. The rate of energy transfer per unit volume between warm dust particles and gas atoms may be obtained from (62), and this heating therefore requires an additional term to be added to Λ_{req} of order

$$\Lambda_{req,dust} \sim 10^{-46}\, T^{1/2}\,(T_d - T) \quad W\, m^3 \qquad (79)$$

Here the same parameters have been used as in (63) except that n_d/n_H has now been taken to be $\sim 10^{-11}$. (79) has a maximum at $T = \frac{1}{3} T_d$, given by

$$\Lambda_{req,dust,max} \sim \frac{2}{3\sqrt{3}} \times 10^{-46}\, T_d^{3/2} \qquad (80)$$

An upper limit to the value of T_d may be estimated from (71) to be on the order of 50 K (where $\mathcal{R}$ has been assumed to be ~ 1 pc), so $\Lambda_{req,dust,max} \approx 10^{-44}$ W m^3. This is much less than the cooling function of an atomic gas at temperatures $T \gtrsim 10$ K, so it is concluded that gas-dust collisions do not significantly affect the equilibrium temperature provided that the gas remains in atomic form.

V-8C Temperature of a mainly molecular disc

The cooling rate for a molecular gas is assumed to be given by (61) and (63):

$$L_{CO} \sim \begin{cases} 6 \times 10^{-46}\, T^2\, n_{H_2}^2 & W\,m^{-3} \quad (n_{H_2} \leq 10^9\, m^{-3}) \\ 2 \times 10^{-28}\, T^3 & W\,m^{-3} \quad (n_{H_2} \gg 10^9\, m^{-3}) \end{cases}$$

$$L_{g-d} \sim 2{\cdot}1 \times 10^{-46}\, n_{H_2}^2\, T^{1/2} (T - T_d) \quad W\,m^{-3}$$

In the low-density regime ($n_{H_2} \leqslant 10^9\ m^{-3}$) cooling due to gas-dust collisions is always much less than that due to CO-emission, since the ratio L_{CO}/L_{g-d} is at least $3\, T^{\frac{1}{2}}$. In this density regime the cooling function can therefore be written in the form

$$\Lambda \sim \Lambda_{CO} \sim 6 \times 10^{-46}\, T^2 \quad W\,m^3 \qquad (81)$$

The magnitude of this cooling function is much less than that of an atomic system at the same temperature, and since the <u>required</u> cooling is slightly larger in the molecular case (see equations (75)) it may be expected that equilibrium temperatures obtained by equating Λ_{CO} and Λ_{req} will be slightly higher than those previously calculated. Equating (81) with (75) therefore gives

$$T_{A1} \sim 140 \left[a \left(\frac{\alpha_* \mathfrak{M}_n}{10^{-2} M_\odot y^{-1}} \right) \left(\frac{R_d}{500\,pc} \right) \right]^{2/3} \left(\frac{10^7 M_\odot}{\mathfrak{M}_d} \right)^{4/3} \left(\frac{\varpi}{100\,pc} \right)^{4/3} \frac{1}{\left(\cos^{-1}(\varpi/R_d)\right)^{2/3}} \quad K \qquad (82\text{-}A1)$$

$$T_{A2} \sim 670 \left[a \left(\frac{\alpha_* \mathfrak{M}_n}{10^{-2} M_\odot y^{-1}} \right) \left(\frac{R_d}{500\,pc} \right)^2 \right] \left(\frac{10^7 M_\odot}{\mathfrak{M}_d} \right)^{3} \left(\frac{\varpi}{100\,pc} \right)^{2} \frac{1}{\left(\cos^{-1}(\varpi/R_d)\right)^{2}} \quad K \qquad (82\text{-}A2)$$

$$T_{B1} \sim 130\left[a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,pc}\right)^{1/2}\right]^{2/3}\left(\frac{10^7 M_\odot}{\mathcal{M}_d}\right)^{4/3}\left(\frac{\varpi}{100\,pc}\right)^{5/3}\frac{1}{\left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^{2/3}}\ K \quad (82\text{-}B1)$$

$$T_{B2} \sim 330\left[a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,pc}\right)\right]\left(\frac{10^7 M_\odot}{\mathcal{M}_d}\right)^{3}\left(\frac{\varpi}{100\,pc}\right)^{3}\frac{1}{\left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^{2}}\ K \quad (82\text{-}B2)$$

from which the temperature may be determined for appropriate combinations of M_d and ϖ such that $n_{H_2} \leq 10^9\ m^{-3}$.

In the central high density regions of the disc, the particle density will be much larger than $10^9\ m^{-3}$, so the appropriate cooling function to use is

$$\Lambda \sim \frac{L_{CO}\left(n_{H_2} \gg 10^9 m^{-3}\right)}{n_{H_2}^2} + \Lambda_{g\text{-}d} \quad (83)$$

Inspection of (61) and (63) shows that the ratio L_{CO}/L_{g-d} is a strong function of radius, varying for T assumed constant $\propto \varpi^{4-6}$ for the various cases. Physically this means that the transition from cooling dominated by dust to that dominated by CO-emission occurs over a small range of radii; consequently a good approximation to the actual temperature distribution will be to consider $\Lambda \sim \Lambda_{CO}$ and $\Lambda \sim \Lambda_{g\text{-}d}$ separately. In the case of CO-dominated cooling equating (83) with (75) implies

$$\left(\frac{\varpi}{100\,pc}\right)_{A1} \sim 370\left(\cos^{-1}(\varpi/R_d)\right)^{1/2} T^{-7/4}\left[a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2} M_\odot y^{-1}}\right)\left(\frac{500\,pc}{R_d}\right)\right]^{1/2} \quad (84\text{-}A1)$$

$$\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)_{A2} \sim 1.75\times10^{-6}\, T^4\left(\frac{\varpi}{100\,pc}\right)^2\frac{1}{\left(\cos^{-1}(\varpi/R_d)\right)^2}\frac{1}{\left[a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2} M_\odot y^{-1}}\right)\cdot\left(\frac{500\,pc}{R_d}\right)\right]} \quad (84\text{-}A2)$$

$$\left(\frac{\varpi}{100\,\mathrm{pc}}\right)_{B1} \sim 120\left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^{2/5} T^{-7/5}\left[a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2}M_\odot y^{-1}}\right)\left(\frac{500\,\mathrm{pc}}{R_d}\right)^{1/2}\right]^{2/5} \qquad (84-B1)$$

$$\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)_{B2} \sim 2{\cdot}25\times10^{-6}\, T^4\left(\frac{\varpi}{100\,\mathrm{pc}}\right)^3 \frac{1}{\left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^2}\,\frac{1}{\left[a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2}M_\odot y^{-1}}\right)\left(\frac{500\,\mathrm{pc}}{R_d}\right)\right]} \qquad (84-B2)$$

whereas in the opposite extreme (cooling dominated by gas-dust collisions) we obtain

$$\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)_{A1} \sim \left[4800\, a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2}M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,\mathrm{pc}}\right)\frac{1}{\cos^{-1}(\varpi/R_d)}\left(\frac{\varpi}{100\,\mathrm{pc}}\right)^2\frac{1}{(T-T_d)}\right]^{1/2} \qquad (85-A1)$$

$$\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)_{A2} \sim \left[1900\, a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2}M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,\mathrm{pc}}\right)^2\frac{1}{\left(\cos^{-1}(\varpi/R_d)\right)^2}\left(\frac{\varpi}{100\,\mathrm{pc}}\right)^2\frac{T^{1/2}}{(T-T_d)}\right]^{1/3} \qquad (85-A2)$$

$$\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)_{B1} \sim \left[4300\, a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2}M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,\mathrm{pc}}\right)\frac{1}{\cos^{-1}\left(\sqrt{\varpi/R_d}\right)}\left(\frac{\varpi}{100\,\mathrm{pc}}\right)^{5/2}\frac{1}{(T-T_d)}\right]^{1/2} \qquad (85-B1)$$

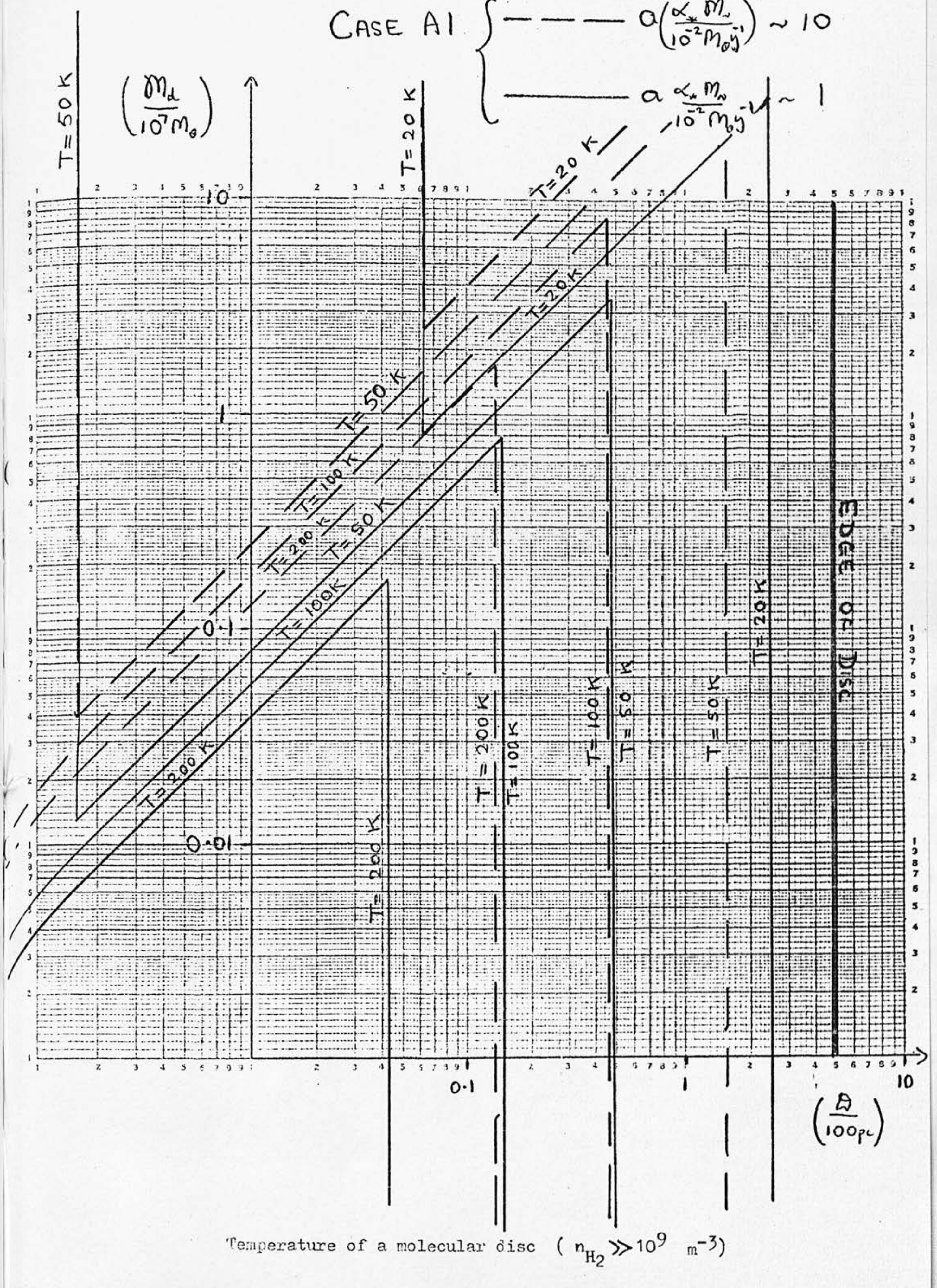
$$\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)_{B2} \sim \left[950\, a\left(\frac{\alpha_* \mathcal{M}_n}{10^{-2}M_\odot y^{-1}}\right)\left(\frac{R_d}{500\,\mathrm{pc}}\right)\frac{1}{\left(\cos^{-1}\left(\sqrt{\varpi/R_d}\right)\right)^2}\left(\frac{\varpi}{100\,\mathrm{pc}}\right)^3\frac{T^{1/2}}{(T-T_d)}\right]^{1/3} \qquad (85-B2)$$

These equations, since they are only valid provided $n_{H_2} \gg 10^9\ \mathrm{m}^{-3}$, are most likely to apply close to the disc centre where $\varpi \ll R_d$. It is therefore a good approximation to substitute a constant of order $\pi/2$ wherever $\cos^{-1}(\varpi/R_d)$ and $\cos^{-1}(\sqrt{\varpi/R_d})$ appear in these formulae, and inspection of (84) then shows that lines of constant temperature are straight lines in the $(\log M_d - \log\varpi)$-plane.

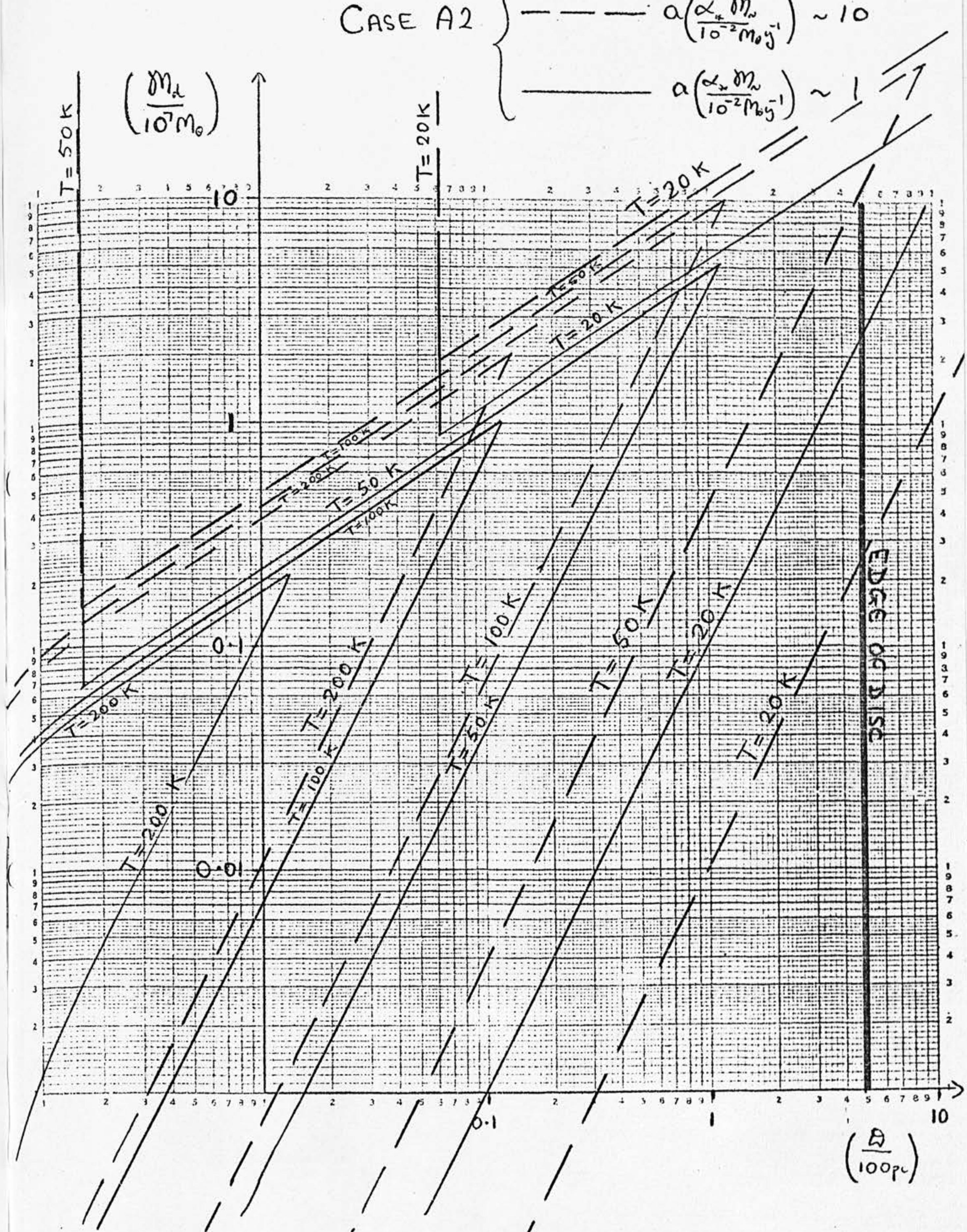
Equation (85) is more complicated, owing to the appearance of the factor $(T - T_d)$ on the RHS, but at radii and temperatures such that $T \gtrsim$ few x T_d, lines of constant T will be nearly straight lines too. At radii where T begins to approach T_d, loci of constant T will begin to bend upwards, and at $T = T_d(\varpi)$ the indicated disc mass will tend to infinity. If the radial variation of dust temperature is assumed to be

$$T_d \sim 10 \left(\frac{100\,pc}{\varpi}\right)^{1/4} \quad K$$

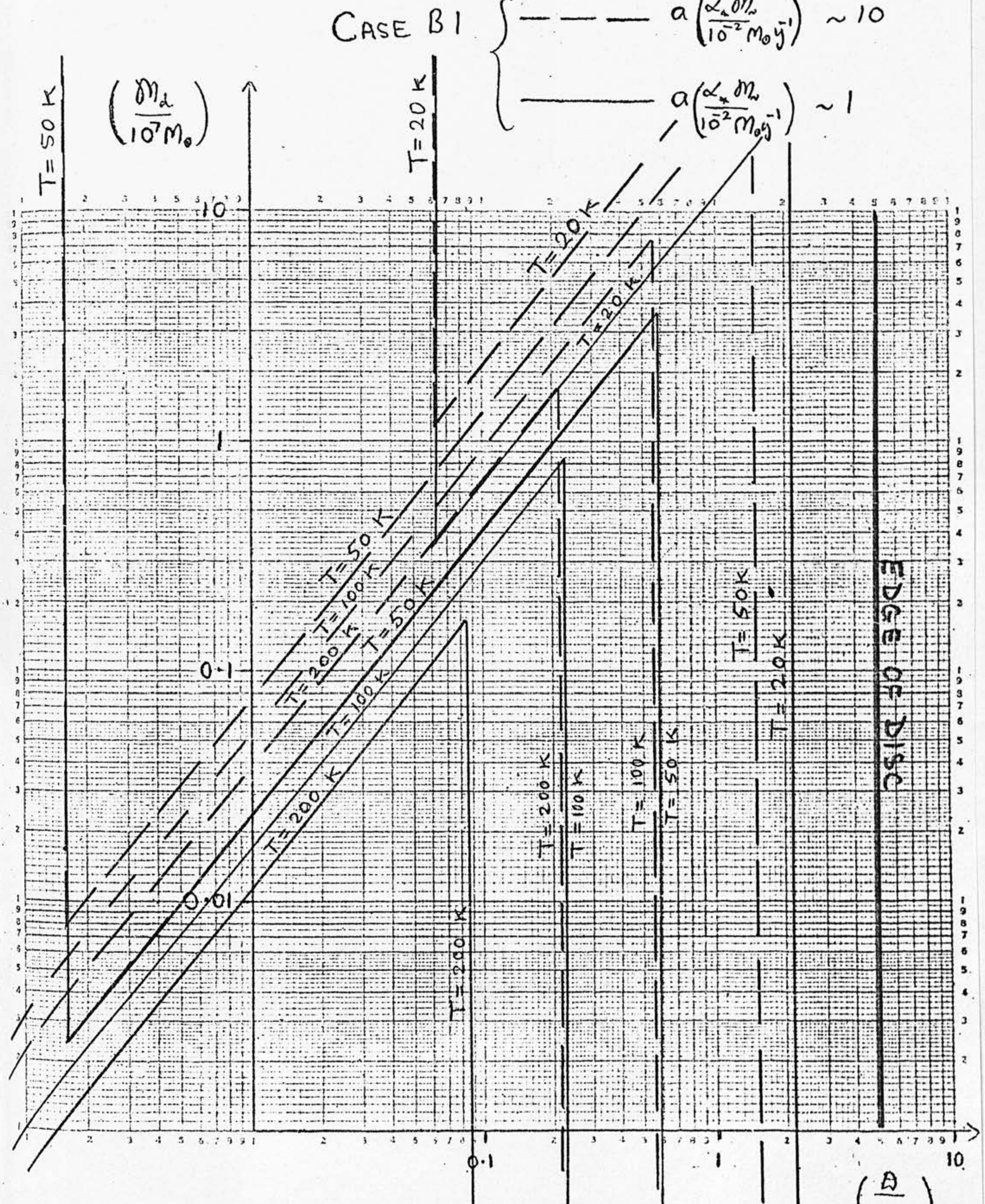
(cf. equation (74)), loci of constant $T \sim T_d$ become straight lines parallel to the M_d-axis at $(\varpi/100\ pc) \sim (10/T)^4$. Combining these results allows the loci of constant T in the $(\log M_d - \log \varpi)$-plane to be represented for illustrative purposes as a combination of three straight lines, which schematically have the form  . These loci are plotted on the graphs overleaf for the same examples that were considered in the case of atomic cooling. For clarity R_d has again been taken to be 500 pc.



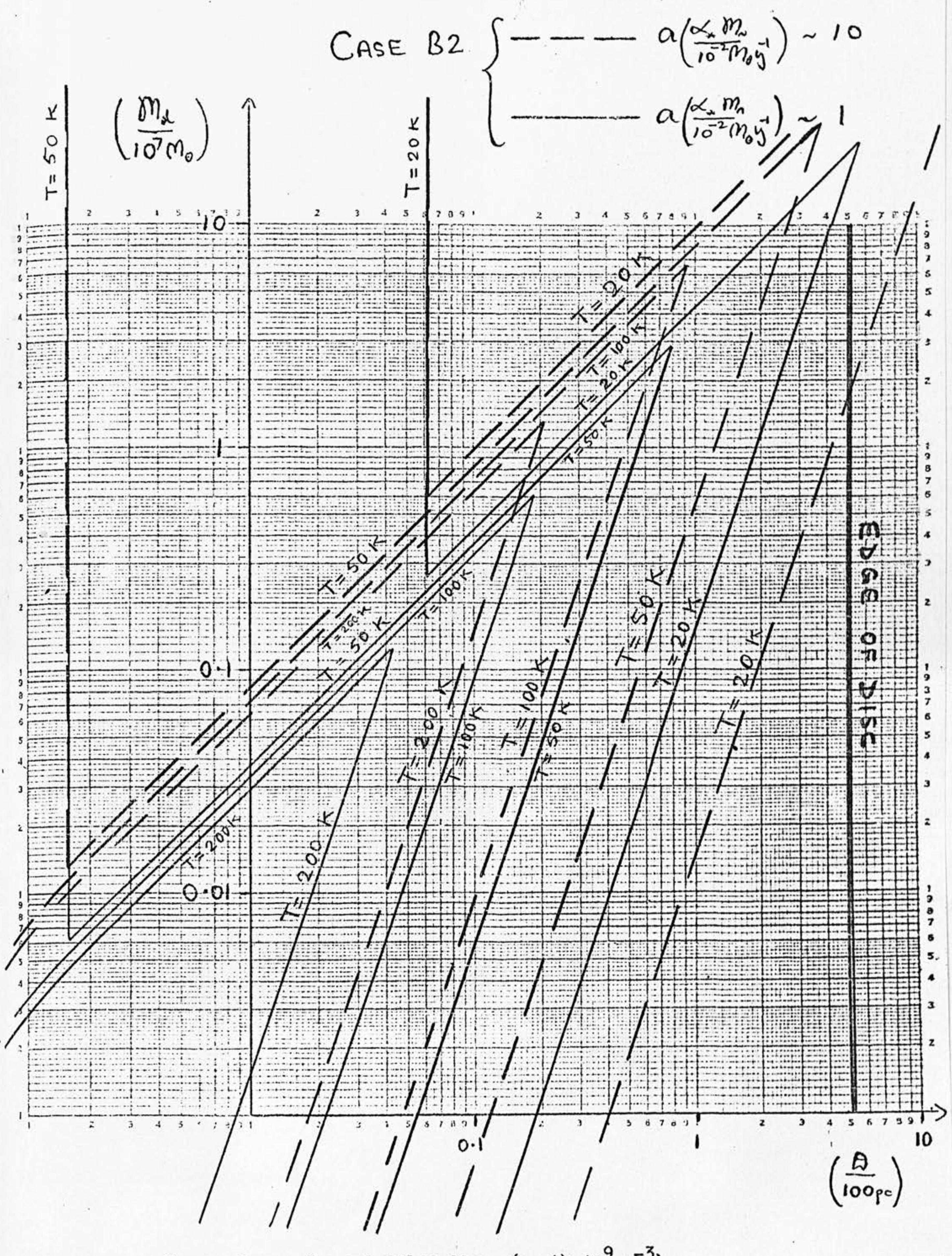
Temperature of a molecular disc ($n_{H_2} \gg 10^9$ m^{-3})



Temperature of a molecular disc ($n_{H_2} \gg 10^9\ m^{-3}$)



Temperature of a molecular disc $(n_{H_2} \gg 10^9\ m^{-3})$



Temperature of a molecular disc $(n_{H_2} \gg 10^9\ m^{-3})$

V-8D Discussion

An important characteristic of the disc which affects most noticeably its detailed physical structure is the rapid density increase towards the centre. This arose essentially from the assumed star density ($\rho_*(r) \propto r^{-2}$), which caused the projected disc density and also the strength of the gravitational attraction towards mid-plane to both increase inwards, thereby leading to $\rho \propto \varpi^{-k}$ where k was in the range 2 - 3 for the various cases discussed. Heating and cooling processes both depend sensitively on the gas density and the physical state of the disc, and - as has been illustrated in the graphs above - this leads to a radial temperature distribution that has a complicated dependence on radius. In order to present the main conclusions of this section, a result mentioned earlier - that Case A discs are typically much more massive than Case B ones when they first become gravitationally unstable - is used. (This result is proved in the next chapter). Results are therefore described for two particular examples: (1) a Case A disc of mass $\sim 10^7$ $M_\odot$; and (2) a Case B disc of mass $\sim 10^5$ $M_\odot$. It is assumed that neither of these discs is close to the onset of gravitational instability, so the discussion will refer only to those graphs concerning Cases A1 and B1 respectively.

(1) Case A1 disc of mass $M_d \sim 10^7$ $M_\odot$

Inspection of the graphs for Case A1 assuming that the disc is in predominantly atomic form (page 98) shows that the expected temperature of a 10^7 $M_\odot$ disc is $\lesssim 20$ K within the central ~ 100 pc. This low value of T, combined with a particle density of order $10^9(100\ \mathrm{pc}/\varpi)^2$ m^{-3} (equation (41)), almost certainly indicates that the gas in this part of the disc will be mainly in molecular form. Reference to the appropriate graph now shows that the central temperature is governed mainly by gas-dust cooling, which is so efficient for $\varpi \leqslant$ few pc that T is kept close to T_d. It should be noted here that the curves drawn should not be extrapolated to radii $\lesssim 1$ pc, because in this region the assumed star density must break down. In a realistic model $\rho_*(\varpi \lesssim 1 \text{ pc}) \sim$ constant, and instead of diverging at the origin roughly as $\varpi^{-1/4}$, T_d may be expected to tend towards a constant value in the central region. The peak value $T_d(0)$ will then be not much different from its value at $\varpi \sim 1$ pc. As ϖ increases, the gas temperature initially decreases outwards in line with the dust temperature ($\propto \varpi^{-1/4}$), but at a radius governed by both M_d and the assumed energy deposited by infalling gas, a point is soon reached where gas-dust collisions can only cool the gas provided that $T \gg T_d$. For the two examples graphed (corresponding to the cases of hot and cold infall respectively) this radius is given by $\varpi \sim 3$ pc and 8 pc. Throughout the region where dust-gas cooling is dominant $\Delta T \equiv T - T_d$ is proportional to ϖ^2, so beyond the radii just mentioned T begins to undergo a rapid increase outwards, $\propto \varpi^2$. This rapid temperature increase continues out to the point at which CO-cooling begins to become dominant (at $\varpi \sim 10$ pc and ~ 20 pc for the cases of hot and cold infall respectively), after which the temperature again decreases outwards, $\propto \varpi^{-4/7}$ (equation (84-A1)). At greater radii still (on the order of ~ 100 pc for a $10^7 \, M_\odot$ disc) cooling due to CO-emission changes to its low-density form ($n_{H_2} \leqslant 10^9 \text{ m}^{-3}$), and the temperature may once again increase outwards, $\propto \varpi^{4/3}$ (equation (82-A1)). At still larger radii it is likely that the disc will begin to contain an increasing fraction of atomic cooling constituents. Since atomic cooling is so much more efficient than CO-emission, the temperature may be expected to fall, only to increase again in the outer regions of the disc, where the effects of increasing cooling efficiency (due to the greater abundance of atomic cooling constituents) are more than offset by the decreasing density there.

(2) Case B1 disc of mass $M_d \sim 10^5 M_\odot$

As in the case just described, inspection of the equilibrium temperatures assuming atomic cooling indicates that the region within $\varpi \sim 10$ pc (depending on the infall rate) might contain a significant fraction of molecular species. On the assumption however that the gas is in mainly atomic form, the temperature of the disc onto which hot material is falling (with $x = 10^{-3}$) is in the range 10 - 20 K at $\varpi \lesssim 1$ pc. This increases rapidly outwards to the limiting value of order 10^5 K at a radius of only ~ 50 pc, indicating that $x \sim 10^{-3}$ is unlikely to be a good approximation except in the very central regions. In the opposite extreme graphed (that in which the infall is relatively cold and $x \sim 0{\cdot}1$) the temperature at radius $\varpi \sim 1$ pc is $\lesssim 10$ K, and at larger radii it rises slowly to a value of order 50 K at $\varpi \sim 50$ pc. Beyond 50 pc the atomic gas temperature increases very rapidly outwards, and a value close to the limiting temperature ($\sim 10^5$ K) is reached at the edge of the disc.

If the disc contains a significant fraction of molecules close to the centre, cooling will be less efficient and equilibrium temperatures somewhat higher. In the case of cold infall, inspection of the appropriate graph shows that in the central regions of a low-mass molecular disc cooling is still dominated by gas-dust collisions (as it was in the example discussed under (1) above); but the region where $T \sim T_d$ has not been reached. Instead, in the case of cold infall, $T \gg T_d$, and the temperature increases outwards $\propto \varpi^{5/2}$ (equation (85-B1)). The temperature at $\varpi \sim 1$ pc is on the order of 200 K. The case of hot infall is slightly different, as dust-gas cooling is not quite dominant at $\varpi \sim 1$ pc. In this case cooling is dominated by CO-emission at ~ 1 pc, leading to $T(\varpi = 1\ \text{pc}) \sim 1800$ K, with $T(\varpi) \propto \varpi^{-5/7}$ (equation (84-B1)). In the case of cold infall CO-cooling does however become important quite close to the centre, and for $\varpi \gtrsim$ few pc $T(\varpi)$ is also $\propto \varpi^{-5/7}$. From these high molecular temperatures it is concluded that a realistic low-mass disc will probably contain a mixture of atomic and molecular cooling constituents, and it is expected that the actual equilibrium temperatures achieved will lie somewhere between the two extremes just described.

A possibility not considered in this work is that the estimate of the heating effect due to infall may be greatly in error. This might occur for example if the infalling gas led only to local heating of the surface layers of the disc so that heating of the bulk of the disc material was constrained to be by conductive and / or radiative processes. In the case in which heating of the main body of the disc is by radiative processes, most of the gas will be heated by collisions with warm dust grains (which absorb the radiation), and an estimate of $\Delta T = T_d - T$ may be obtained by equating L_{CO} with L_{g-d}. This gives:

$$\Delta T \sim \begin{cases} 3T^{3/2} & n_{H_2} \leq 10^9\ \text{m}^{-3} \\ \left(\dfrac{10^9\ \text{m}^{-3}}{n_{H_2}}\right)^2 T^{5/2} & n_{H_2} \gg 10^9\ \text{m}^{-3} \end{cases} \qquad (86)$$

In both these cases the gas temperature must be less than that of the dust (since dust is now responsible for heating the gas), and since this is not affected significantly by inclusion of Γ_1 in the expression for thermal balance of a grain, it is concluded that a disc in which the energy of infalling material is dissipated radiatively at the surface will also be very cold, with a mean temperature $\lesssim 30$ K assuming T_d given by (74).

In conclusion, the most important result to emerge from this preliminary study of the detailed disc structure is that typical discs more massive that $\sim 10^5 \, M_\odot$ can absorb even hot infalling material without heating up. Nuclear discs can therefore behave as 'sinks' for stellar mass loss, and outflowing galactic winds are not likely to occur in the nuclear regions of galaxies. If recurrent nuclear activity is to be explained on the present model it is necessary that an active (Seyfert) phase of evolution does not remove all the gas from the nuclear interstellar medium, and amounts of order $10^5 \, M_\odot$ (depending on the details of the model) must remain behind to act as the 'seed' for the growth by stellar mass loss of a new disc. Observations both of active Seyfert galaxies (e.g. Walker, 1968) and of possibly post-active ones such as our own Galaxy (see Chapter VII) suggest that this requirement is indeed met. Thus, if the evolution of one disc can be shown to lead to a period of nuclear activity then recurrence follows. The gas density in the model discs increased inwards very rapidly, and the central temperatures were found to be very low, indicating that much of the gas would probably be in molecular form. If discs in elliptical galaxies are also predominantly in molecular form, this might partly explain their low apparent gas masses (inferred from H I measurements) which are usually explained by appeal to a gas-removal process such as a galactic wind (e.g. Faber & Gallagher, 1976).

VI EVOLUTION OF THE DISC AND IMPLICATIONS FOR THEORIES OF NUCLEAR ACTIVITY

In the theory described in this thesis, the evolution of the nuclear disc can be divided into two main phases: (1) an extended period of slow mass-growth during which viscous evolution may be neglected; and (2) the onset of a star-forming gravitational instability followed by the formation of a massive non-stellar object, presumed to be the ultimate cause of a period of nuclear activity. This chapter presents the results of a preliminary investigation into these two phases of the disc's evolution, and ends (Section VI-3) with a discussion of the chief implications that the present considerations have for particular theories of nuclear activity.

VI-1 Evolution up to the onset of gravitational instability

VI-1A Review of the problem

An important feature of the present theory is that stellar mass loss flows into the central regions of galaxies where it forms a dense gaseous disc which stores the gas for long periods between successive cycles of activity. This places stringent requirements on the model, and in order for storage to be effective during the disc's growth phase, it is necessary that the viscous timescale be long compared with other timescales of interest (for example the mean interval between active periods). In the present model the viscous timescale is given by equation (23), which, for a V_c = constant rotation law, reduces to

$$\tau_{visc} \sim 8\cdot 6 \times 10^6 \left(\frac{\varpi}{1\,\mathrm{pc}}\right)\left(\frac{10\ \mathrm{km\,s^{-1}}}{\bar{c}}\right)^2 \quad y \qquad (24)$$

On the assumption that the mean turbulent velocity $\bar{c}$ is on the same order as the sound speed (since strongly supersonic motions are quickly damped) this may be re-written in the form

$$\tau_{visc} \sim \sqrt{2}\ \frac{\bar{m}}{kT}\ \varpi V_0$$

$$\Longrightarrow \tau_{visc} \sim 4\cdot 6 \times 10^9 \left(\frac{\varpi}{1\,\mathrm{pc}}\right)\left(\frac{10\,\mathrm{K}}{T}\right) \quad . y \qquad (87)$$

where it has been assumed that $\bar{m} = 1\cdot 3\ m_H$, the value appropriate for an atomic disc. It should be emphasised that this equation is not expected to apply within the central $\sim$1 pc of the disc, because this is the region within which the assumed star distribution becomes unrealistic. An improved estimate of τ_{visc} in this region requires an improved model of the central star density. If $\rho_*(r)$ is assumed to tend towards a constant value within $\varpi \sim 1$ pc (cf. Sanders & Lowinger, 1972), the circular velocity will tend towards uniform rotation, leading to a somewhat larger estimate of τ_{visc} (e.g. equation (23)). It should be noted however that even in this case the circular velocity very close to the centre ($\varpi \ll 1$ pc) might not be exactly uniform, because at very small radii the disc mass within ϖ may well become comparable with the stellar mass within the same radius. For a model in which $\rho_* \sim$ constant within a central core of radius $R_c \sim 1$ pc, it may be shown that the fractional disc mass within ϖ for $\varpi \ll 1$ pc is small whichever assumed rotation law is chosen (i.e. whether Case A or Case B). Because of this, provided that $\bar{c}$ remains small, effects due to possible viscous evolution very close to the disc centre are assumed to be negligible. At radii $\varpi \gtrsim 1$ pc, (87) does give a reasonable estimate of τ_{visc}, and because it has been shown that even low-mass discs are probably very cold in their central regions, viscous flows may indeed be neglected (to a first approximation) during the disc's slow growth phase. It is of interest to express this result in terms of the expected viscous flux across radius ϖ. Using Lynden-Bell & Pringle's (1974) result for a steady-state disc, the inward flux across ϖ is given (equation (22)) by:

$$F = 2\pi \nu \sigma \left(1 - \frac{\varpi}{V_c}\frac{dV_c}{d\varpi}\right) \qquad (22)$$

This reduces for $V_c(\varpi)$ assumed constant to

$$F_A \sim \frac{1}{\sqrt{2}} \frac{kT}{\bar{m}} \frac{\mathcal{M}_d}{R_d} \frac{1}{V_0} \cos^{-1}(\varpi/R_d) \qquad (88\text{-}A)$$

$$F_B \sim \frac{1}{2\sqrt{2}} \frac{kT}{\bar{m}} \frac{\mathcal{M}_d}{(\varpi R_d)^{1/2}} \frac{1}{V_0} \cos^{-1}\left(\sqrt{\varpi/R_d}\right) \qquad (88\text{-}B)$$

which implies (again assuming $\bar{m} = 1{\cdot}3\ m_H$)

$$F_A \sim 4{\cdot}4 \times 10^{-5} \left(\frac{T}{100K}\right)\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)\left(\frac{500pc}{R_d}\right).\cos^{-1}\left(\varpi/R_d\right) \quad M_\odot y^{-1} \qquad (89\text{-}A)$$

$$F_B \sim 4{\cdot}9 \times 10^{-5} \left(\frac{T}{100K}\right)\left(\frac{\mathcal{M}_d}{10^7 M_\odot}\right)\left(\frac{500pc}{R_d}\right)^{1/2}\left(\frac{100pc}{\varpi}\right)^{1/2}.\cos^{-1}\left(\sqrt{\varpi/R_d}\right) \quad M_\odot y^{-1} \qquad (89\text{-}B)$$

These radial flows are very much less than the expected infall rate, so the approximation of no significant viscous evolution is well founded.

The disc grows slowly in mass until eventually its density becomes sufficiently high for the onset of gravitational instability. The exact value of σ_{crit} depends weakly on details of the particular model calculated (Goldreich & Lynden-Bell, 1965a,b), but an estimate of sufficient accuracy for the present purpose (correct to within a factor of about 2) is given by (29):

$$\sigma_{crit} \sim \frac{\kappa V_s}{\pi G} \qquad (29)$$

Since $\kappa = \sqrt{2}\,\frac{V_c}{\varpi}\left(1 + \frac{\varpi}{V_c}\frac{dV_c}{d\varpi}\right)^{1/2}$, in the present model this reduces to

$$\sigma_{crit} \sim \frac{\sqrt{2}}{\pi} \frac{V_0 V_s}{\varpi G} \sim \frac{\sqrt{2}}{\pi} \frac{\mathcal{M}_n}{\varpi R_n} v_s \qquad (90)$$

where the dimensionless parameter $v_s = V_s/V_0$ has again been introduced, and use been made of the relation $V_0^2 = (GM_n/R_n)$. The surface densities of Case A and Case B discs are given by (19), so the critical disc mass, $\mathcal{M}_{crit}(\varpi)$, at which instability first occurs at radius ϖ is

$$\frac{\mathcal{M}_{crit,A}}{\mathcal{M}_n} \sim 2\sqrt{2}\ \frac{R_d}{R_n}\ v_s(\varpi) \Big/ \cos^{-1}(\varpi/R_d) \qquad (91\text{-}A)$$

$$\frac{\mathcal{M}_{crit,B}}{\mathcal{M}_n} \sim 4\sqrt{2}\ \frac{(R_d\varpi)^{1/2}}{R_n}\ v_s(\varpi) \Big/ \cos^{-1}\left(\sqrt{\varpi/R_d}\right) \qquad (91\text{-}B)$$

The discussion in the previous chapter showed that although the actual radial variation of temperature in a particular disc was in general quite a complicated function, the overall trend (due mainly to the outward decrease in density) was one in which T increased outwards. The r.m.s. gas velocity is proportional to $T^{\frac{1}{2}}$ (equation (40)), so this quantity takes its smallest values close to the disc centre. The phase of slow mass-growth of the disc ends with the first onset of gravitational instability at some radius, and inspection of equations (91) shows that in the present model both Case A and Case B discs first become unstable close to the origin (i.e. $\varpi_{crit} \lesssim 1$ pc; the equations are not generally valid at smaller radii). For Case A the critical disc mass at which instability first occurs can be estimated by setting $\varpi \lesssim 1$ pc and assuming that the disc is predominantly in molecular form with a central temperature $T(0) \sim T_d(0) \sim 30$ K. This procedure, together with the previously assumed values of M_n, R_n and R_d ($10^{10}\ M_\odot$, 1 kpc and 500 pc respectively) give the approximate result

$$\mathcal{M}_{crit,A} \sim 3{\cdot}3 \times 10^7\ M_\odot \qquad (92\text{-}A)$$

The critical Case A mass is not very sensitive to the precise value of $\varpi_{crit} \lesssim 1$ pc, but it should be noted that the critical radius may depend on the exact details of the temperature distribution. Equation (91-B) shows that $M_{crit,B} \approx \varpi^{\frac{1}{2}}_{crit}$, so in this case a realistic choice of $\varpi_{crit} \lesssim 1$ pc is more important. Using the same values for M_n, R_n and R_d as above, and assuming that the Case B disc is in mainly atomic form with a central temperature of order 20 K thus implies

$$\mathcal{M}_{crit,B} \approx 10^6 \left(\frac{\varpi_{crit}}{0{\cdot}1\,pc}\right)^{1/2} M_\odot \qquad (92\text{-}B)$$

Even for a representative critical radius as large as 0·1 pc Case B discs are typically less massive than Case A ones. This qualitative conclusion is confirmed by the more detailed estimates made below (Section VI-1D). In both disc types gravitational instability is found to occur first close to the origin ($\varpi \lesssim 1$ pc), and because the assumed star distribution breaks down in this central region, more detailed estimates must be based on a more realistic star density. An improved star distribution is outlined in the next section, and results obtained by its use are described for each case in the succeeding sections; VI-1C and VI-1D respectively.

VI-1B Outline of an improved model

An improved star distribution not suffering from the disadvantage of diverging at the origin may be defined by

$$\rho_*(r) = \begin{cases} \dfrac{M_n'}{4\pi R_n} \dfrac{1}{R_c^2} & 0 \leq r \leq R_c \\[2ex] \dfrac{M_n'}{4\pi R_n} \dfrac{1}{r^2} & R_c \leq r \leq R_n \end{cases} \tag{93}$$

where M_n' is related to the nuclear mass M_n by

$$M_n = M_n' \left(1 - \frac{2}{3}\frac{R_c}{R_n}\right) \tag{94}$$

As before, parameters which roughly reproduce observations of our Galaxy are $M_n \sim 10^{10}\ M_\odot$ and $R_n \sim 1$ kpc. The radius of the central core R_c is taken to be $R_c \sim 1$ pc.

The circular velocity as a function of radius is

$$V_c(\varpi) = \begin{cases} V_0' \dfrac{\varpi}{\sqrt{3}\,R_c} & 0 \leq \varpi \leq R_c \\[2ex] V_0' \left(1 - \dfrac{2}{3}\dfrac{R_c}{\varpi}\right)^{1/2} & R_c \leq \varpi \leq R_n \end{cases} \tag{95}$$

where V_0' is defined by

$$V_0 = V_0' \left(1 - \frac{2}{3}\frac{R_c}{R_n}\right)^{1/2} . \tag{96}$$

The rotation within the central region ($\varpi \lesssim R_c \sim 1$ pc) is now uniform, and the circular velocity increases rapidly at larger radii towards the limiting value V_0.

The gravitational potential ϕ_* is

$$\phi_*(r) = \begin{cases} -V_0'^2 \left[\ln\left(\frac{R_n}{R_c}\right) + \frac{1}{2} - \frac{1}{6}\left(\frac{r}{R_c}\right)^2\right] & 0 \leq r \leq R_c \\ -V_0'^2 \left[1 + \ln\left(\frac{R_n}{r}\right) - \frac{2}{3}\left(\frac{R_c}{r}\right)\right] & R_c \leq r \leq R_n \end{cases} \tag{97}$$

Following the same procedure as that in Section V-4, the disc density neglecting self-gravity is

$$\ln\left(\frac{\rho(\varpi,z)}{\rho(\varpi,0)}\right) = -\frac{3}{2V_s^2}\left[\phi(\varpi,z) - \phi(\varpi,0)\right] \tag{98}$$

which implies

$$\ln\left(\frac{\rho(\varpi,z)}{\rho(\varpi,0)}\right) \sim \begin{cases} -\frac{3}{2}\frac{V_0'^2}{V_s^2}\cdot\frac{1}{3}\frac{z^2}{R_c^2} & 0 \leq \varpi < R_c \\ -\frac{3}{2}\frac{V_0'^2}{V_s^2}\left[\ln\left(1+\frac{z^2}{\varpi^2}\right) - \frac{4}{3}\frac{R_c}{\varpi}\left(1 - \frac{1}{\sqrt{1+z^2/\varpi^2}}\right)\right] & R_c \leq \varpi \leq R_n \end{cases} \tag{99}$$

The disc width is therefore

$$H(\varpi) \sim \begin{cases} \sqrt{2\pi}\,\frac{V_s}{V_0'}\cdot R_c & 0 \leq \varpi < R_c \\ \sqrt{\frac{2\pi}{3}}\,\frac{V_s}{V_0'}\left(1 - \frac{2}{3}\frac{R_c}{\varpi}\right)^{-1/2}\cdot\varpi & R_c \leq \varpi \leq R_n \end{cases} \tag{100}$$

where, as before, it has been assumed that $(V_0'/V_s)^2 \gg 1$, so that a good approximation to (99) is obtained by considering only terms to first order in z^2/ϖ^2.

In the case that the z-structure of the disc is dominated by self-gravity, as it is when the disc first becomes gravitationally unstable, the density distribution $\rho(\varpi,z)/\rho(\varpi,0)$ is independent of the assumed star density. $H(\varpi)$ is then given by (28); i.e.

$$H(\varpi) = \frac{2}{3\pi}\frac{V_s^2}{G\sigma} \qquad (28)$$

The critical disc width (i.e. $H(\varpi)$ when $\sigma = \sigma_{crit}$) is therefore

$$H_{crit}(\varpi_{crit}) = \frac{2}{3}\frac{V_s}{K} \qquad (101)$$

$$\Rightarrow H_{crit}(\varpi_{crit} \leq R_c) \sim \frac{1}{\sqrt{3}}\frac{V_s}{V_0'} R_c \sim 3\times10^{-3}\left(\frac{V_s}{1\ \mathrm{km\,s^{-1}}}\right)\left(\frac{R_c}{1\ \mathrm{pc}}\right)\ \mathrm{pc} \qquad (102)$$

where the result $K(\varpi \leq R_c) = \frac{2}{\sqrt{3}}\frac{V_0'}{R_c}$ (derived from (95)) has been used.

The critical disc density may be estimated to be on the order of

$$\rho_{crit} \sim \sigma_{crit}/H_{crit} \sim 3K^2/2\pi G$$

$$\Rightarrow \rho_{crit}(\varpi_{crit} \leq R_c) \sim \frac{2}{\pi}\frac{\mathcal{M}_n'}{R_n R_c^2} \qquad (103)$$

Adopting the usual values for M_n, R_n and R_c thus implies

$$\rho_{crit}(\varpi_{crit} \leq R_c) \sim 4\cdot3\times10^{-13}\ \mathrm{kg\ m^{-3}} . \qquad (104)$$

which corresponds to a hydrogen number density of order $2 \times 10^{14}\ \mathrm{m^{-3}}$.

A consequence of the model's more realistic circular velocity (now decreasing to zero at the origin) is that a more complicated definition of Case A is required. Previously this Case was defined by the requirement that the net rotational velocity $V_{rot}(\varpi)$ was constant; however because physically it is necessary that $V_{rot} \leq V_c$, close to the origin a slightly different definition must be considered. For the purpose of the present investigation Case A was defined to have a distribution of angular momentum characterised by

$$v_{rot,A} = b\left(1 - e^{-x/c_A}\right) \tag{105-A}$$

where the dimensionless quantities $x = \varpi/R_n$ and $v_{rot} = V_{rot}/V_o$ have been introduced. Although it is not strictly necessary, an analogous re-definition of Case B was also considered:

$$v_{rot,B} = b\, x\, e^{-x/c_B} \tag{105-B}$$

The constants in these formulae were chosen to be

$$b = 0{\cdot}5\ ; \quad c_A = 0{\cdot}001\ ; \quad c_B = 100 \tag{106}$$

so that the outer radius of the disc would still be close to 500 pc, and the two cases would correspond very closely to the Cases A and B discussed previously.

With these modifications, the development of the theory follows exactly the same lines as before, except that all the formulae become much more complicated. An analytic determination of the equilibrium surface densities is no longer practicable, so the two cases were computed numerically and the ratio σ/σ_{crit} determined for various assumptions as to the radial variation of V_s. The purpose of these computions was two-fold: First, to obtain an improved estimate of the Case B critical disc mass, which depends on the disc's surface density within the central parsec; and secondly, by considering different forms for the radial variation of V_s, to try and determine how sensitive ϖ_{crit} and M_{crit} are to this factor. In particular, since it was found that the first gravitational instability occurred in both cases close to the origin, it is important to consider the possibility that one effect of this might be a purely local heating at the disc centre which might temporarily remove the instability and perhaps allow the next one to develop away from the centre. The 'improved' star distribution (93), although allowing estimates of ϖ_{crit} and M_{crit} to be made within the region $\varpi \lesssim 1$ pc, is still only a rough approximation to a realistic model, and consequently results applying to the central region are not expected physically to be very accurate. For this reason, although computations were made with various values of the constants b and c, a detailed discussion of the various cases computed was not thought worthwhile. A brief summary of the most important results is given in the next two sections.

VI-1C Results for Case A

The analytic results derived earlier showed that the expression for the disc mass at which gravitational instability occurs at radius ϖ was given by (91). For $R_d = 500$ pc, $R_n = 1$ kpc and $M_n = 10^{10}\ M_\odot$ this reduces to

$$\mathcal{M}_{crit,A}(\varpi) \sim \frac{\sqrt{2}\ V_s(\varpi)}{\cos^{-1}(\varpi/R_d)} \cdot 10^{10}\ M_\odot \qquad (107)$$

The disc first becomes gravitationally unstable when the RHS of (107) takes its minimum value, and, as was noted above, for cases in which V_s is constant or increases outwards, this condition occurs at the origin. A molecular disc with central temperature $T(0) \sim T_d(0) \sim 30$ K therefore becomes unstable when its mass reaches $M_{crit} \sim 3{\cdot}3 \times 10^7\ M_\odot$. Reference to the graphs showing the temperature distribution of a self-gravitating molecular disc with mass $\gtrsim 10^7\ M_\odot$ (Case A2, page 109), shows that close to the centre (within ≈ 10 pc, depending on the heating rate) $T \sim T_d(\varpi)$. At larger radii T increases rapidly with increasing radius and then decreases again, when CO-cooling becomes dominant. The lowest temperatures reached in this outer region are no lower than that achieved by gas-dust cooling at $\varpi \sim 10$ pc (i.e. $T(\varpi \gtrsim 50\text{ pc}) \gtrsim 20$ K), so because the denominator on the RHS of (107) is larger at smaller radii it may be concluded that instability will occur first at the smaller radius. Evaluation of (107) at $\varpi \sim 10$ pc with $T \sim 20$ K shows that inclusion of the radial temperature distribution of $V_s(\varpi)$ $(\propto T^{\frac{1}{2}} \lesssim \varpi^{-1/8})$ in the central region has the effect of lowering the critical mass to $M_{crit} \sim 2{\cdot}7 \times 10^7\ M_\odot$. Thus, in the analytic approximation, the disc first becomes gravitationally unstable at $\varpi \sim 10$ pc. It may be verified that this conclusion is unaltered even if atomic cooling processes should lower the temperature in the outer regions to $T \sim 10$ K, although the critical mass for $\varpi \sim 50$ pc and an atomic disc is only slightly larger ($\sim 2{\cdot}8 \times 10^7\ M_\odot$) than the value just calculated.

The sensitivity of ϖ_{crit} to the details of the assumed temperature distribution arises in this model essentially because equation (107) with V_s = constant has a very broad weakly defined minimum, allowing small changes in V_s to significantly affect its exact location. The numerical calculation, using the improved star distribution and the slightly different assumed distribution of net rotational velocity, led to a final equilibrium surface density whose maximum no longer occurred exactly at the origin. This was due to the different interplay between V_{rot} and V_c within $R_c \sim 1$ pc, and led in turn to a somewhat stronger minimum in the calculated function corresponding to the RHS of (107). For V_s = constant, the critical radius was found to be $\varpi_{crit} \sim 0{\cdot}7$ pc, and assuming (as before) a mainly molecular disc with a central temperature of order 30 K, the critical mass was $M_{crit} \sim 2{\cdot}3 \times 10^7\ M_\odot$. At radii $\varpi \gtrsim$ few pc the numerical results were hardly distinguishable from the analytic ones (as indeed they should be, since the two models approach each other in this region), but because the new critical mass ($\sim 2{\cdot}3 \times 10^7\ M_\odot$) was less than that obtained at $\varpi \sim 10$ pc with $T \sim 20$ K (i.e. $2{\cdot}7 \times 10^7\ M_\odot$) inclusion of the more realistic form of V_s did not this time affect the position of ϖ_{crit}

From these results it is concluded that the Case A disc first becomes gravitationally unstable when its mass reaches a critical value on the order of $3 \times 10^7\ M_\odot$. The precise radius at which instability first occurs, ϖ_{crit}, is difficult to determine exactly, since a large portion of the disc becomes liable to gravitational instability at approximately the same time. It is probable that ϖ_{crit} is as much determined by 'random' effects (such as infall of a particularly dense cloud) as it is by specific details of the model. However, with this proviso in mind, the analytic approximation and the more detailed numerical model both led to the result that the initial instability would occur close to the galactic centre ($\varpi \lesssim 10$ pc and $\lesssim 1$ pc respectively). Any viscous evolution (here neglected) would increase the surface density close to the centre, thereby strengthening the above tendency. For these reasons it is concluded that the Case A disc will first become unstable at $\varpi_{crit} \lesssim 10$ pc.

At this point of the disc's evolution the unstable region will fragment into individual clouds of characteristic mass m_c, depending on σ_{crit} and the wavelength λ_{crit} of the fastest-growing unstable mode. Goldreich & Lynden-Bell (1965a,b) showed that both σ_{crit} and λ_{crit} depended in detail on particular features of the assumed model, and it has already been mentioned that the adopted criterion for instability (equation (25)) is only accurate to within a factor of order 2. In what follows it will be assumed that λ_{crit} is given approximately by

$$\lambda_{crit} \sim 2\pi T \qquad (108)$$

where T is the disc thickness, defined (following Goldreich & Lynden-Bell) by

$$T = \sigma / \bar{\rho}$$

and as before

$$\bar{\rho} = \int \rho^2 dz \Big/ \int \rho \, dz$$

The characteristic mass of an unstable fragment is

$$m_c \sim \pi \left(\frac{\lambda_{crit}}{2}\right)^2 \sigma_{crit} \qquad (109)$$

and since in the present model $\top \sim \frac{1}{\pi} . V_s^2 . \frac{1}{G\sigma}$ (cf. equation (28)) this reduces to

$$m_c \sim \pi^2 V_s^3 \Big/ K(\varpi_{crit})\, G \qquad (110)$$

Now $K \doteq \sqrt{2}\, \frac{V_c}{\varpi}\left(1 + \frac{dV_c}{d\varpi}\frac{\varpi}{V_c}\right)^{1/2}$, so the improved model star distribution implies the approximate result

$$K(\varpi) \sim \begin{cases} \frac{2}{\sqrt{3}}\, \frac{V_0}{R_c} & 0 \lesssim \varpi \lesssim R_c \\ \sqrt{2}\, \frac{V_0}{\varpi} & R_c \lesssim \varpi \lesssim R_n \end{cases} \qquad (111)$$

Thus (110) becomes

$$\left(\frac{m_c}{\mathcal{M}_n}\right) \approx \begin{cases} \frac{\sqrt{3}\,\pi^2}{2}\, \upsilon_s^3 \left(\frac{R_c}{R_n}\right) & 0 \lesssim \varpi_{crit} \lesssim R_c \\ \frac{\pi^2}{\sqrt{2}}\, \upsilon_s^3 \left(\frac{\varpi_{crit}}{R_n}\right) & R_c \lesssim \varpi_{crit} \lesssim R_n \end{cases} \qquad (112)$$

Values of m_c determined from (112) should be treated with some care, since the assumed values of σ_{crit} and λ_{crit} are both uncertain by factors of order two (depending on specific details of the model such as the vertical structure of the disc, the character of the differential rotation and the nature of the perturbation that is applied to the system). The final expression is thus possibly uncertain by a factor as large as 10, and estimates of fragment masses obtained by its use should be treated only as indicative. On the assumption that the disc is in mainly molecular form with a central temperature of order 30 K, adopting $\varpi_{crit} \sim R_c \sim 1$ pc gives the result

$$\mathcal{M}_c \approx 3 \left(\stackrel{\times}{\div} 10\right) \ \mathcal{M}_\odot \tag{113}$$

The first objects to form in the unstable Case A disc are thus likely to be ordinary stars. If the instability first occurs at a significant distance from the centre ($\varpi_{crit} \sim 100$ pc, say), (112) then implies that initial fragment masses are probably too large to form single stars. In this case (by analogy with star formation in the outer parts of galaxies) it seems likely that each initial unstable fragment will ultimately collapse to form a star cluster or stellar association.

VI-1D Results for Case B

In contrast to the Case A disc (in which M_{crit} is quite well defined but Θ_{crit} is uncertain), the analytic result for Case B showed that gravitational instability of the disc was likely to occur first at the origin (i.e. $\Theta_{crit} = 0$), when the disc mass was very small (strictly zero). The minimum at the origin in the analytic expression for $M_{crit}(\Theta)$ is strong and defined very clearly (due mainly to the rapid increase in σ_B as $\Theta \rightarrow 0$), so small variations of V_s from a constant value will not greatly affect its position. However because the assumed star distribution breaks down within the central parsec, an improved calculation is now necessary in order to obtain a realistic estimate of M_{crit}. The result of this calculation confirms the qualitative conclusion arrived at earlier, that typical Case B discs are likely to be less massive than their Case A counterparts, and for a disc with a central r.m.s. velocity dispersion V_s the critical mass is found to be

$$\mathcal{M}_{crit,B} \sim 4{\cdot}8 \times 10^{4} \left(\frac{V_s}{1\ \mathrm{km\,s^{-1}}}\right) \quad \mathcal{M}_\odot \tag{114}$$

Evaluation of the appropriate value of V_s is more difficult for the (low-mass) Case B disc than the (massive) Case A one, since the central temperature depends sensitively on the assumptions made as to the heating and cooling rates (i.e. whether or not hot infall, and whether or not atomic cooling). Inspection of the appropriate graphs shows that central temperatures lie in the approximate ranges $10 \lesssim T \lesssim 1000$ K and $200 \lesssim T \lesssim 500$ K for atomic and molecular discs respectively, indicating that V_s lies in the broad range

$0{\cdot}6 \lesssim (V_s/1\ \mathrm{km\ s^{-1}}) \lesssim 6$ and $2 \lesssim (V_s/1\ \mathrm{km\ s^{-1}}) \lesssim 3$ for the same two cases. From this result it is concluded that the critical Case B mass is given approximately by

$$\mathcal{M}_{crit,B} \approx 10^5 \left(\stackrel{\times}{\div} 3\right) \quad \mathcal{M}_\odot \qquad (115)$$

and that instability will occur first at the centre.

The mass of the first unstable fragments may be estimated in the same way as for the Case A disc, and by use of (112) the result

$$m_c \sim 8{\cdot}5\ v_s^3 \left(\frac{R_c}{R_n}\right) \mathcal{M}_n$$

is obtained. Adopting the usual values of R_c, R_n and M_n thus implies

$$m_c \approx 10 \left(\frac{V_s}{1\ \mathrm{km\ s^{-1}}}\right)^3 \quad \mathcal{M}_\odot$$

where the same uncertainty as before $(\stackrel{\times}{\div} 10)$ is implied. It is thus reasonable to assume initially that stars will form in the central regions of this disc too, although until the central velocity dispersion is much better defined the possibility that individual clouds as massive as $\sim 10^4\ M_\odot$ might form must be considered. Clouds this massive would comprise a significant fraction (~ 10 %) of the whole disc mass (almost the whole disc within $\varpi \sim 1$ pc), so the later evolution of this kind of disc would be very complicated.

VI-2 Conditions immediately after the onset of gravitational instability

In the previous section it has been shown firstly that both Case A and Case B discs first become subject to gravitational instability close to the centre, and secondly that the first objects to form are likely to be ordinary stars. Although the disc might conceivably always produce low-mass stars, this possibility does not seem very realistic, and it is therefore assumed that a few newly-formed stars will be massive enough to ionise their immediate surroundings and produce an H II region (cf. equations (113) and (116)). The radius of this region may be estimated from the radius of the corresponding Stromgren sphere, and for stars of spectral class B0 to O5 it can be shown (Allen, 1973; p.267) that

$$R_s \sim (1-5) \times 10^{-4} \left(\frac{n}{10^{14}\ m^{-3}}\right)^{-2/3} \quad pc \qquad (117)$$

A reasonable order of magnitude estimate of the critical disc density is $n \sim 10^{14}\ m^{-3}$ (equation (104)), and it is therefore concluded, by comparing (117) with (102), that H II regions produced by newly-formed massive stars will have diameters much less than the disc width $H(\varpi)$.

Shock and ionisation fronts produced by a first generation of massive stars will propagate into the neutral material surrounding the first gravitationally unstable region and lead to the fast development of a cold dense layer in which further massive star formation will occur (Elmegreen & Lada, 1977). An estimate of the timescale for formation of a second generation of stars in this region (which separates the outwardly moving shock and ionisation fronts) is given by Elmegreen & Lada (their equation (18)) by:

$$t_I \sim 2.3 \times 10^6 \left(\frac{n}{10^9 m^{-3}}\right)^{-1/2} y$$

In the present problem (in which $n \sim 10^{14}\ m^{-3}$) new stars will be formed very soon after the birth of the first massive stars. Since the Kelvin-Helmholtz contraction timescale for O-stars is $\lesssim 4 \times 10^4$ y (e.g. Elmegreen & Lada, 1977), the initial (star-forming) gravitational instability could lead to several generations of massive stars on a timescale $\sim$ few x 10^5 years.

The theoretical problem of determining the later evolution of the system is arguably one of the most difficult in astrophysics, involving as it does (1) star formation, (2) the evolution of molecular clouds and H II regions, (3) viscosity in the interstellar medium, and (4) either accretion onto a black hole or the formation and evolution of some other kind of massive object, such as a spinar or supermassive star. Each one of these factors involves large uncertainties and unknowns, and for this reason it is impractical to attempt a comprehensive and detailed discussion of all the various possibilities. Instead, the principal aims of the remainder of this chapter, in line with the spirit of the general enquiry, are first to outline a sequence of events which might plausibly lead to the formation of an active nucleus, and secondly to use these preliminary results to argue in favour of a particular model of the active nucleus, thereby isolating an important theoretical problem for future investigation.

VI-2A Stability of the disc against disruption by star formation

The onset of massive star formation close to the disc centre will lead to a greatly increased local heating rate, and it is important to show that this is insufficient to lead to the disc's disruption. This section presents two very rough estimates of the mean temperature of a uniform gaseous disc in which massive star formation is taking place, and then considers in somewhat greater detail the more realistic case of a non-uniform disc.

First, if it is assumed that a fraction f of the disc material fragments into stars, and that each star radiates energy at less than the Eddington limit (i.e. $L_* \lesssim L_E \sim 1{\cdot}4 \times 10^{31}\ M/M_\odot$ W), then cooling at $T \sim 10^4$ K can dominate heating provided that $f \lesssim 0{\cdot}99$. This result follows by considering the balance between heating and cooling as represented by

$$\Lambda(T)\left[(1-f)n\right]^2 \lesssim 1{\cdot}4 \times 10^{31} \cdot \frac{f n \bar{m}}{M_\odot} \quad W$$

and where $\Lambda(T \sim 10^4\ \mathrm{K}) \approx 10^{-35}\ \mathrm{W\ m^3}$ and $\rho = n\bar{m}$ has been taken to be approximately equal to $1/20\,\rho_{crit}$, implying $n \sim 10^{13}\ \mathrm{m^{-3}}$. This density would arise if random gas motions ('turbulence') driven by massive star formation had a r.m.s. velocity dispersion $\bar{c} \sim 20\,V_{s,crit} \approx 20\ \mathrm{km\ s^{-1}}$, where $V_{s,crit}$ (the r.m.s. velocity dispersion just before gravitational instability) is assumed to be of order $1\ \mathrm{km\ s^{-1}}$. If f is of order $0{\cdot}5$ this expression becomes

$$\Lambda_{req}(T) \lesssim 3 \times 10^{-39}\ \mathrm{W\ m^3}$$

which indicates (see the table on page 79) equilibrium temperatures in the range 50 - 1000 K, depending on the fractional ionisation.

This result is only valid provided that the disc is optically thin, and although the opacities of typical interstellar gases at temperatures $T \lesssim$ few x 10^3 K are indeed very small in the absence of dust (e.g. Cox & Guili, 1968; p.353), it has been shown above that the probable presence of dust will ensure that the disc has a large optical depth at visual wavelengths. If the disc should be opaque to cooling radiations, a better estimate of its equilibrium temperature may be obtained by equating the luminosity produced by newly-formed stars to the energy radiated from the disc surface. Again using the Eddington limit as an upper limit to the power radiated by stellar sources, we obtain

$$2\pi R^2 \sigma T^4 \sim 1{\cdot}4 \times 10^{31} \frac{f\, \mathcal{M}_d(\varpi \leq R)}{M_\odot} \quad W$$

The disc mass within $\varpi \sim R_c \sim 1$ pc is $\lesssim 5 \times 10^4\ M_\odot$ (see Section VI-2B below), so it may be concluded that $T \lesssim 250$ K. Even if the same luminosity is assumed to be concentrated within a spherical region of radius such that the circular velocity equals $V_s \sim 20$ km s^{-1}, the required temperature is still $\lesssim 500$ K.

These results show, provided that the gas is distributed fairly uniformly, that a burst of massive star formation occurring in the central most dense regions of the disc is unable to raise its mean temperature sufficiently to cause its destruction. In reality the assumption of uniformity is unlikely to be valid, as shocks caused by supernovae, expanding H II regions and early-type stellar winds will all tend to create density inhomogeneities and drive large-scale mass-motions. In order to show that even these phenomena do not lead to excessively large systematic motions of cool material, we consider by way of example the effect of a powerful early-type stellar wind. Castor et al (1975) have discussed this phenomenon in some detail, and note that winds associated with observed early-type stars (spectral types earlier than B2) have typical mass loss rates $\dot{M}_w \sim 10^{-6}\ M_\odot\ y^{-1}$ and terminal velocities $V_w \sim 2000\ \mathrm{km\ s^{-1}}$. The radius of the shell of swept-up matter at time t is given by their equation (6); i.e.

$$R_s(t) \sim 0.76 \left(\frac{\frac{1}{2}\dot{m}_w V_w^2}{\rho_0}\right)^{1/5} t^{3/5} \qquad (118)$$

where ρ_0 is the density of the ambient medium. The velocity of the expanding shell is

$$\dot{R}_s(t) \sim \left(\tfrac{3}{5} \times 0.76\right) \cdot \left(\frac{\frac{1}{2}\dot{m}_w V_w^2}{\rho_0}\right)^{1/5} t^{-2/5} \qquad (119)$$

Under the conditions of extremely high density appropriate to the present problem, the shell very quickly enters a 'snowplough' phase in which rapid cooling of swept-up gas occurs (see equation (10) of Castor et al, 1975). This cold shell continues its outward expansion according to (118) until part of it eventually breaks out of the disc. At this point the hot interior gas is expected to flow rapidly out of the burst 'bubble' leaving the remainder of the shell to continue its motion in the plane of the disc. (It should be noted however that because the shell is no longer driven by the stellar wind, after this point has been reached R_s will no longer be governed by (118)). In order to obtain a rough estimate as to the likely size of the mass-motions generated by wind-driven shells in the disc, it is assumed firstly that the moment of break-out t_B can be estimated to order of magnitude by equating $R_s(t_B)$ to half the disc's semi-width; i.e.

$$R_s(t_B) \sim \frac{1}{4} H(\varpi) \qquad (120)$$

and secondly that the r.m.s. velocity dispersion in the disc is given approximately by $V_s \sim \dot{R}_s(t_B)$. The disc width within the central region ($\varpi \lesssim R_c$) is given by (100), so combining this with (120) and (118) gives

$$0.76\left(\frac{\frac{1}{2}\dot{m}_\omega V_\omega^2}{\rho_0}\right)^{1/5} t_B^{3/5} \sim \frac{1}{2}\sqrt{\frac{\pi}{2}}\,\frac{V_s}{V_0}\,R_c$$

(119) implies

$$\dot{R}_s(t_B) \times t_B \sim \left(\frac{3}{5}\times 0.76\right)\cdot\left(\frac{\frac{1}{2}\dot{m}_\omega V_\omega^2}{\rho_0}\right)^{1/5} t_B^{3/5}$$

$$\Rightarrow V_s \times t_B \sim \frac{3}{5}\cdot\frac{1}{2}\sqrt{\frac{\pi}{2}}\,\frac{V_s}{V_0}\,R_c$$

$$\text{i.e} \qquad t_B \sim \frac{3}{10}\sqrt{\frac{\pi}{2}}\,\frac{R_c}{V_0} \qquad (121)$$

where the relation $V_s = \dot{R}_s(t_B)$ has been used. Substituting this value of t_B into (119) and making use of the relation

$$\rho_0 \sim \rho_{crit} \times \frac{V_{s,crit}}{V_s} \sim \frac{2}{\pi}\,\frac{m_n}{R_n R_c^2}\left(\frac{V_{s,crit}}{V_s}\right)$$

implies

$$V_s \sim \left(\frac{3}{5}\times 0.76\right)^{5/4}\cdot\left(\frac{50}{9}\cdot\dot{m}_\omega V_\omega^2\cdot\frac{G}{V_{s,crit}}\right)^{1/4}$$

$$\Longrightarrow V_s \sim 6.5\left(\frac{\dot{m}_\omega}{10^{-6} M_\odot y^{-1}}\right)^{1/4}\left(\frac{V_\omega}{2000 \text{ km s}^{-1}}\right)^{1/2}\left(\frac{1 \text{ km s}^{-1}}{V_{s,crit}}\right)^{1/4} \text{ km s}^{-1} \qquad (122)$$

From this result it is concluded that random mass-motions in the disc are unlikely to much exceed several tens of km s^{-1}, and that they will certainly be insufficient to lead to the total destruction of the system.

VI-2B Formation and mass of a proto-nucleus

The burst of star formation is expected to change the disc from its previously quiescent state characterised by $V_s \sim (V_{sound}^2 + \bar{c}^2)^{\frac{1}{2}} \sim 2^{\frac{1}{2}} V_{sound}$, to a strongly turbulent state in which the gas pressure is dominated by the random mass-motions with $\bar{c} \gg V_{sound}$. Present understanding of the structure and dynamics of the interstellar medium does not allow an accurate estimate of $\bar{c}$ to be made, but observations of our Galaxy do indicate that cloud motions with $\bar{c} \gg V_{sound}$ can occur. In particular, observations of our Galaxy's H I 'nuclear disc' feature (Sanders & Wrixon, 1973; Cohen & Davies, 1976) indicate the presence of random motions with velocities up to of order 50 km s^{-1}. In what follows it will be assumed that a representative value of $\bar{c}$ is $\sim 20 \text{ km s}^{-1}$.

Viscous friction caused by the large random gas motions will transport matter into the centre at a rate given approximately by (22), i.e.

$$F \sim 2\pi \nu \sigma$$

Substituting for ν and setting $\sigma = \sigma_{crit}$, this becomes

$$F_{crit} \sim \frac{2}{3} \cdot \frac{\bar{c}^2 V_{s,crit}}{G} \qquad (123)$$

This reduces to

$$F_{crit} \sim 0.06 \left(\frac{\bar{c}}{20 \text{ km s}^{-1}}\right)^2 \left(\frac{V_{s,crit}}{1 \text{ km s}^{-1}}\right) M_\odot \, y^{-1} \qquad (124)$$

If the galactic nucleus is assumed to contain a massive black hole, the luminosity generated by disc accretion at the above rate may be estimated from (13); i.e.

$$L \sim 10^{38} \left(\frac{\eta}{0.3}\right)\left(\frac{\bar{c}}{20 \text{ km s}^{-1}}\right)^{2}\left(\frac{V_{s,crit}}{1 \text{ km s}^{-1}}\right) \quad W \tag{125}$$

which is very close to the luminosity of a typical Seyfert galaxy (e.g. Weedman, 1976a). The present theory is thus not at variance with the black hole model, although it is argued later that other theoretical models of the nature of nuclear activity are more successful at explaining the whole range of observational data. The important conclusion to be drawn from (125) is that the present theory <u>can</u> lead to luminosities comparable with those of observed Seyfert nuclei.

In general, unless black holes are assumed to be primordial in origin (e.g. Ryan, 1972), the first cycle of activity must lead to the initial formation of a spinar or supermassive star. On the black hole hypothesis this object is assumed to undergo eventual gravitational collapse (e.g. Lynden-Bell, 1969), but on other theories (e.g. Fricke, 1973, 1974) it is possible that the body will end its evolution by explosive disruption. In this case, nuclear activity must be understood as due to evolution of the massive object formed at the centre of the disc, and second and subsequent cycles of activity have essentially the same cause as the first. It is therefore of great interest to enquire as to the type of object most likely to be formed by the rapidly evolving disc, and to ask whether or not the evolution of this body does lead to the formation of a black hole.

An upper limit to the mass of any such object formed from the disc is given by $M \sim M_{crit}$. For Cases A and B respectively this is $\sim 3 \times 10^{7}\ M_{\odot}$ and $\sim 10^{5}\ M_{\odot}$. Bodies of smaller mass will be formed if, as seems likely, not all the disc can flow into the centre at a fast enough rate; a better estimate of the limiting mass being given by that contained within a region of radius defined by $\tau_{visc} \lesssim \tau_{evol}$. It is unlikely that τ_{evol} will be much greater than $\sim 10^6$ y, so using this value and τ_{visc} determined from (24), the radius out to which viscous effects might be important in the formation of a massive central object is

$$\varpi_{visc} \lesssim 0.5\left(\frac{\bar{c}}{20\ \mathrm{km\,s^{-1}}}\right)^2\left(\frac{\tau_{evol}}{10^6\ \mathrm{y}}\right) \quad \mathrm{pc}$$

In order to obtain a rough indication as to the limiting masses obtained by this argument, the fractional disc mass contained within ϖ was calculated using equilibrium surface densities determined from the improved model star distribution (equation (93)). The results of this calculation are given in the table below, where as usual it has been assumed that $R_n \sim 1$ kpc and $R_d \sim 500$ pc.

Radius ϖ (pc)	Fractional disc mass within ϖ	
	Case A	Case B
0·1	10^{-5}	0·005
1	0·002	0·05
10	0·03	0·2
100	0·3	0·6

Inspection of the above formula shows that ϖ_{visc} probably lies in the range $1 \lesssim (\varpi_{visc}/1\ \mathrm{pc}) \lesssim 10$, indicating that the mass of the central object formed at the origin will be in the range

$5 \times 10^4 \lesssim (M/M_\odot) \lesssim 10^6$ for Case A, and $5 \times 10^3 \lesssim (M/M_\odot) \lesssim 2 \times 10^4$ for Case B.

An alternative estimate of the mass of the first supermassive object to form after the occurrence of the initial (star-forming) gravitational instability, may be obtained by assuming that the gas flowing inwards by viscous processes accumulates at the centre, with large velocity dispersion, until it too becomes gravitationally unstable. Mestel (1965) gives the Jeans mass of a spherical cloud supported by random gas motions with magnitude $\bar{c}$ as

$$\mathcal{M}_J \sim \frac{3.8\ \bar{c}^3}{G^{3/2}} \left(\frac{3}{4\pi\rho}\right)^{1/2}$$

If it is assumed that $\mathcal{M}_J \sim \frac{4}{3}\pi\rho R^3$ this reduces to

$$\mathcal{M}_J \sim (3.8)^{2/3}\ \bar{c}^2 R/G$$

The circular velocity in the central region $\varpi \lesssim 1$ pc is given by (95), i.e.

$$V_c(\varpi \lesssim 1\,\mathrm{pc}) \sim \frac{\varpi}{\sqrt{3}R_c} V_o'$$

and if it is assumed that the accumulating gas extends over a region of radius R, defined by $R \sim \varpi(V_c = \bar{c})$, we obtain

$$\left(\frac{\mathcal{M}_J}{\mathcal{M}_n}\right) \sim (3.8)^{2/3} \sqrt{3}\ v_s^3 \left(\frac{R_c}{R_n}\right)$$

where $v_s = V_s/V_o \sim \bar{c}/V_o$.

The critical mass obtained in this way is roughly half that given by the approximation used in the earlier discussion of the fragmentation of the disc (equation (112)). Adopting the same values of R_n, R_c and M_n that were used above thus gives

$$\mathcal{M}_J \sim 4 \times 10^4 \left(\frac{\bar{c}}{20 \text{ km s}^{-1}}\right)^3 \mathcal{M}_\odot \qquad (126)$$

which is in good agreement with the limiting masses obtained above by the viscous timescale argument.

This estimate may be improved by direct use of the results of Goldreich & Lynden-Bell (1965a) which relate specifically to the properties of uniformly rotating discs, and are therefore applicable to the present problem. Instead of (25), we now have

$$\frac{\pi G \bar{\rho}_{crit}}{4\Omega^2} \sim 0.73 \qquad (127)$$

which implies (cf. equation (103))

$$\rho_{crit} \sim (0.73)\,\frac{2}{\pi}\,\frac{\mathcal{M}_n}{R_n R_c^2} \sim 3.1 \times 10^{-13} \text{ kg m}^{-3} \qquad (128)$$

The critical surface density is now

$$\sigma_{crit} \sim (0.73)^{1/2}\,\frac{\kappa V_s}{\pi G}$$

$$\Rightarrow \sigma_{crit} \sim (0.73)^{1/2}\,\frac{2}{\pi\sqrt{3}}\,V_s\,\frac{\mathcal{M}_n}{R_n R_c} \qquad (129)$$

where κ has been taken equal to $2\Omega = \frac{2}{\sqrt{3}}\frac{V_0}{R_c}$ (equation (111)).

The wavenumber of the first gravitationally unstable mode is given by $k\,T \sim 1.4$ (cf. equation (108)), indicating

$$\lambda_{crit} \sim \frac{3\pi}{1.4}\,H_{crit} \qquad (130)$$

where use has been made of the relations $\lambda_{crit} = 2\pi/k$ and $\Gamma = 3/2$ H. The mass of the first unstable fragment is therefore

$$\mathcal{M} \sim \pi \left(\frac{\lambda_{crit}}{2}\right)^2 \sigma_{crit}$$

$$\Rightarrow \mathcal{M} \sim (1{\cdot}4)^{-2} (0{\cdot}73)^{-1/2} \frac{\pi^2 V_s^3}{k G} \qquad (131)$$

$$\text{i.e} \quad \mathcal{M} \sim (1{\cdot}4)^{-2} (0{\cdot}73)^{-1/2} \frac{\pi^2 \sqrt{3}}{2} v_s^3 \left(\frac{R_c}{R_n}\right) \mathcal{M}_n \qquad (132)$$

which should be compared with the more approximate formulae used earlier (equations (110) and (112)). Adopting the usual values of the parameters M_n, R_n and R_c, we have:

$$\mathcal{M} \sim 4{\cdot}6 \times 10^4 \left(\frac{V_s}{20 \text{ km s}^{-1}}\right)^3 M_\odot \qquad (133)$$

which, it should be noted, is also in good agreement with the mass obtained by the argument based on the Jeans mass (equation (126)). The radius of the first self-gravitating fragment to form in the now rapidly evolving disc is R $\sim \lambda_{crit}/2$, so

$$R \sim \frac{3\pi}{2{\cdot}8} H_{crit} \sim \frac{\pi\sqrt{3}}{2{\cdot}8} \frac{1}{(0{\cdot}73)^{1/2}} v_s R_c \qquad (134)$$

$$\Rightarrow R \sim 0{\cdot}22 \left(\frac{V_s}{20 \text{ km s}^{-1}}\right) R_c \qquad (135)$$

This body forms at the centre of a cool rapidly rotating disc, so it too may be expected to be flattened and disc-like in form. This may be verified by considering the ratio of rotational and random kinetic energies; i.e.

$$\frac{E_{rot}}{E_{rand}} \sim \frac{1}{2} I \Omega^2 \Big/ \frac{1}{2} m V^2 \qquad (136)$$

For a uniform disc $I = \frac{1}{2} M R^2$, and since initially Ω is given approximately by $\Omega \sim V_c(\varpi)/\varpi \sim V_o/\sqrt{3}\, R_c$, (136) becomes

$$\frac{E_{rot}}{E_{rand}} \sim 0.86$$

It is thus concluded that unless angular momentum transport during the fragmentation process is particularly efficient, the newly-formed 'proto-nucleus' will initially resemble a uniformly rotating self-gravitating disc.

VI-2C Evolution of the proto-nucleus, and possible structure and evolution of an active nucleus

The active nucleus must form by evolution of the proto-nucleus, whose structure has been shown above to initially closely resemble that of a uniformly rotating self-gravitating disc. The collapse of uniformly rotating gas clouds has been discussed by a number of authors (e.g. Mestel, 1963, 1965; Cameron, 1969a,b), and it has been found generally that for a given distribution of initial angular momentum, two qualitatively different types of self-gravitating disc may be formed. In the first, it is found that the surface density σ depends only weakly on radius and that the angular velocity Ω is uniform; whereas in the second, it is found that the disc is much more axially condensed, and $\Omega \propto \varpi^{-1}$. Which type of disc is actually formed depends on subtle details of the collapse process, although Mestel (e.g. 1965) has emphasised that discs of the former type are liable to collapse into those of the latter type if perturbed sufficiently that the increased degree of central condensation is able to overcome the stabilising effect of the centrifugal force. The reverse evolution does not occur, so it is concluded that in the present case (where strong perturbations are indeed expected) the initially uniformly rotating disc will quickly evolve into the axially condensed type characterised by a strong differential rotation.

The surface density of the V = constant disc may be shown (Mestel, 1963; equation (56)) to be

$$\sigma(\varpi) = \frac{V^2}{2\pi R \varpi}\left[1 - \frac{2}{\pi}\sin^{-1}\varpi/R\right] \qquad (137)$$

where V is the (constant) circular velocity and R the disc's outer radius. Since the disc mass is

$$\mathcal{M} \sim \int_0^R 2\pi \varpi \sigma(\varpi)\, d\varpi$$

we have

$$V^2 = \frac{\pi}{2} \frac{G\mathcal{M}}{R} \tag{138}$$

In a realistic case V must drop to zero at the origin (cf. Cameron, 1969a,b), so very close to the centre the surface density will depart from that given by (137). However, for the purpose of the present investigation, aimed simply at getting rough estimates of the gross properties of the nucleus, this detail wil be ignored. The first 'interesting' object which forms in the galactic nucleus by evolution of the large-scale nuclear disc is therefore a massive self-gravitating differentially rotating disc, with mass and radius given approximately by equations (133) and (135) respectively, and surface density σ defined by

$$\sigma(\varpi) \sim \frac{\mathcal{M}}{4R} \frac{1}{\varpi} \left[1 - \frac{2}{\pi} \sin^{-1}(\varpi/R) \right] \tag{139}$$

Since this body is in strong differential rotation, turbulence driven by hot stars, supernovae, infalling matter and the passage of high-velocity stars from the nuclear star cluster will lead by viscous transport to the rapid formation of an increasingly massive central condensation. It is interesting to note the similarity between the surface density (139) and that expected for a steady-state viscous disc with $\nu \propto \varpi$ (cf. equation (21)). In the present model we have

$$\nu \sim \frac{1}{3\sqrt{2}} \bar{c}^2 \frac{\varpi}{V}$$

so the viscous timescale is given roughly by (23); i.e.

$$\tau_{visc} \sim 3\sqrt{2}\ \frac{\varpi V}{\bar{c}_n^2} \qquad (140)$$

where $\bar{c}_n$ has been introduced to distinguish it from $\bar{c}$, the mean turbulent velocity in the larger non self-gravitating disc surrounding the proto-nucleus.

Now $V = \left(\frac{\pi}{2}\frac{G\mathcal{M}}{R}\right)^{1/2}$, so using $\mathcal{M} = \pi R^2 \sigma_{crit}$ and the definitions of R and σ_{crit} (equations (134) and (129)) we obtain:

$$V \sim \frac{\pi}{\sqrt{2\cdot 8}}\, V_s \qquad (141)$$

$$\Rightarrow \quad \tau_{visc} \sim \frac{3\pi}{\sqrt{1\cdot 4}}\ \frac{\varpi V_s}{\bar{c}_n^2} \qquad (142)$$

Combining this result with the definition of the initial radius R of the disc (equation (134)) gives

$$\tau_{visc} \sim \left(\frac{3}{1\cdot 4}\right)^{3/2} \frac{\pi^2}{2}\, \frac{1}{(0\cdot 73)^{1/2}} \left(\frac{V_s}{\bar{c}_n}\right)\left(\frac{\varpi}{R}\right)\left(\frac{R_c}{V_0}\right) \qquad (143)$$

$$\text{i.e} \quad \tau_{visc} \sim 8\cdot 5\times 10^4 \left(\frac{V_s}{\bar{c}_n}\right)\left(\frac{\varpi}{R}\right)\ \text{y} \qquad (144)$$

where the usual values of R_c and V_0 have been assumed.

The rotation period at ϖ is of order $2\pi\varpi/V$, so

$$\frac{\tau_{visc}}{P} \sim \frac{3}{\sqrt{2}}\,\frac{1}{\pi}\left(\frac{V}{\bar{c}_n}\right)^2$$

$$\text{i.e} \quad \frac{\tau_{visc}}{P} \sim \frac{3\pi}{\sqrt{2}}\,\frac{1}{(2\cdot 8)}\left(\frac{V_s}{\bar{c}_n}\right)^2 \sim 2\cdot 4\left(\frac{V_s}{\bar{c}_n}\right)^2 \qquad (145)$$

From these results it is concluded that viscosity will be important in forming a central condensation.

A rough estimate as to the rate at which viscosity transports matter inwards is given by (22); i.e.

$$F_n \sim 2\pi \nu \sigma$$

$$\Rightarrow \quad F_n \sim \frac{\pi}{6\sqrt{2}} \frac{\bar{c}_n^2 \mathcal{M}}{R V} \tag{146}$$

where use has been made of the relations $\nu = \frac{1}{3}\frac{\bar{c}_n^2}{\kappa} \sim \frac{1}{3\sqrt{2}} \cdot \frac{\bar{c}_n^2 \varpi}{V}$ and $\sigma \sim \mathcal{M}/4R\varpi$. Using (138) thus gives

$$F_n \sim \frac{1}{3\sqrt{2}} \frac{\bar{c}_n^2 V}{G}$$

and since V and V_s are related by (141), this reduces to

$$F_n \sim \frac{\pi}{6\sqrt{1.4}} \frac{\bar{c}_n^2 V_s}{G} \tag{147}$$

$$\text{i.e} \quad F_n \sim 0.8 \left(\frac{\bar{c}_n}{20 \text{ km s}^{-1}}\right)^2 \left(\frac{V_s}{20 \text{ km s}^{-1}}\right) M_\odot \, y^{-1} \tag{148}$$

The inward flux due to viscosity in the surrounding disc is F_{crit}, given by (123); so

$$\frac{F_n}{F_{crit}} \sim \frac{\pi}{4\sqrt{1.4}} \left(\frac{\bar{c}_n}{\bar{c}}\right)^2 \left(\frac{V_s}{V_{s,crit}}\right) \tag{149}$$

Since $V_s \sim 20 \text{ km s}^{-1} \gg V_{s,crit} \sim 1 \text{ km s}^{-1}$, it may be concluded that viscous transport in the proto-nucleus will be at least as effective as that in the surrounding disc.

A detailed theory of the evolution of the proto-nucleus should include the complex and possibly important effects of (1) accretion, leading to mass (M) and angular momentum (J) increasing; (2) equatorial mass loss and non-thermal radiation, leading to M and J decreasing; (3) gravitational contraction, leading to energy generation, and Ω, T and ρ increasing; and (4) nuclear burning, leading to energy generation and (possibly) to total or partial disruption. Until this has been done it is not possible to say with certainty what kind of nucleus will eventually be formed. It does, however, seem likely that in the first instance the system will form a rapidly rotating spinar-like object (cf. Ozernoy & Usov, 1971, 1973; Sturrock & Barnes, 1972) rather than a slowly rotating supermassive star as discussed for example by Fricke (1973, 1974). Since an important factor in favour of spinar models is their ability to produce large non-thermal luminosities (often far exceeding the thermal Eddington limit component), powers comparable to those of the most luminous quasars might in principle be attained by the evolution of even quite low-mass objects. The detailed evolutionary history of a spinar depends to a great extent on the particular model that is investigated: on its mass, mode of formation, rate of energy loss, and on whether it is a 'hot' high-entropy object (Ozernoy & Usov, 1971, 1973; Fowler, 1971) or a 'cold' rapidly rotating disc (e.g. Wagoner, 1971; Salpeter & Wagoner, 1971; Wagoner & Salpeter, 1972; Sturrock & Barnes, 1972; Hara & Ikeuchi, 1976). In the present theory, the arguments above lead to the suggestion that initially the proto-nucleus will resemble a cold rapidly rotating disc, and it might therefore be expected - certainly before nuclear energy generation becomes important - that the nucleus so formed would also initially more closely resemble the low-entropy spinar types than the high-entropy ones. The detailed evolution of disc-shaped models has received less theoretical attention than that of the high-entropy ones, and for this reason and also because the evolution of both kinds of object is qualitatively the same (although there are important quantitative differences; Salpeter & Wagoner, 1971), we here describe the possible structure and evolution of the active nucleus by reference to the work of Ozernoy & Usov (1971, 1973), which was concerned with models of 'hot' spinars.

A spinar radiates both thermally and non-thermally; its thermal power being given approximately by the Eddington relation:

$$L_{th} \sim L_E = \frac{4\pi G c}{\kappa} \mathcal{M} \tag{150}$$

and its non-thermal ('magnetic dipole') power by

$$L_{md} \sim \frac{2}{3c^3} \frac{4\pi}{\mu_0} B_p^2 R^6 \Omega^4 \sin^2 \chi \quad W \tag{151}$$

Here κ is the opacity, and in the second formula the body has been assumed to be rotating uniformly with a poloidal magnetic field of strength B_p (tesla), directed at an angle χ to the rotation axis. Assuming that the opacity is dominated by electron scattering, we have

$$\kappa \sim 0{\cdot}02\,(1+X) \quad m^2\ kg^{-1} \tag{152}$$

and if the spinar is composed of gas with 'normal' cosmic abundances, the hydrogen mass fraction X is $\sim 0{\cdot}73$ (Allen, 1973; p.30). The expression for the thermal power (equation (150)) thus reduces to

$$L_E \sim 1{\cdot}4 \times 10^{31} \left(\frac{\mathcal{M}}{M_\odot}\right) \quad W \tag{153}$$

During its evolution the hot spinar loses energy by radiation, contracts and spins faster until eventually matter begins to be lost from its equator. Its angular velocity is thus given, to within a factor of order unity (depending on precise details of the model such as the shape, density distribution, and strength of the magnetic field; Ozernoy & Usov, 1971), by:

$$\Omega \sim \Omega_{cr} \sim \left(\frac{GM}{R^3}\right)^{1/2} \qquad (154)$$

If the magnetic field within the spinar is constant (equal to B_p), and its density distribution can be approximated by a spherical polytrope of index 3, then the ratio (ζ) of magnetic to gravitational energy is

$$\zeta \equiv \frac{E_m}{|E_{grav}|} \sim \left(\frac{4}{3}\pi R^3 \frac{B_p^2}{2\mu_0}\right) \Big/ \left(\frac{3}{2}\frac{GM^2}{R}\right) \qquad (155)$$

$$\Rightarrow \quad B_p \sim 3\left(\frac{\mu_0 G}{4\pi}\right)^{1/2} \frac{M}{R^2}\,\zeta^{1/2} \quad \mathrm{T} \qquad (156)$$

$$\text{i.e} \quad B_p \sim 1.54\left(\frac{M}{10^6 M_\odot}\right)\left(\frac{10^{14}\,\mathrm{m}}{R}\right)^2 \zeta^{1/2} \quad \mathrm{T} \qquad (157)$$

This corresponds to equation (2.11) of Ozernoy & Usov (1973). Substituting this expression into (151), measuring Ω in units of Ω_{cr} and R in units of the gravitational radius $R_g = 2GM/c^2$, thus gives

$$L_{md} \sim \frac{3}{8}\frac{c^5}{G}\,\zeta\left(\frac{\Omega}{\Omega_{cr}}\right)^4\left(\frac{R_g}{R}\right)^4 \sin^2\chi \quad \mathrm{W} \qquad (158)$$

$$\text{i.e} \quad L_{md} \sim 1.4\times10^{52}\,\zeta\left(\frac{\Omega}{\Omega_{cr}}\right)^4\left(\frac{R_g}{R}\right)^4 \sin^2\chi \quad \mathrm{W} \qquad (159)$$

This corresponds to equation (6.7) of Ozernoy & Usov (1973), which is too small by a factor of 10. (This numerical error recurs in some other formulae in their paper, and is also propagated in later work, e.g. Saslaw, 1974; Ozernoy & Usov, 1977).

The spinar evolves through loss of energy by radiation from one quasi-static equilibrium configuration to another; the 'quasi-static' phase of evolution occurring when the radius is within the limits $R_{cr} \lesssim R \lesssim R_{qs}$. Here

$$R_{qs} \sim 6 \times 10^{14} \left(\frac{z}{0.1}\right)^{2/5} \left(\frac{\mathcal{M}}{10^6 M_\odot}\right)^{3/5} \quad m \qquad (160)$$

and

$$R_{cr} \sim 4 \times 10^{10} \left(\frac{0.1}{z}\right) \left(\frac{\mathcal{M}}{10^6 M_\odot}\right) \quad m \qquad (161)$$

where R_{cr} is the radius below which stability is broken by effects due to General Relativity, and z is the ratio of rotational to gravitational energy. These formulae are derived from equations (6.10) and (6.11) of Ozernoy & Usov (1973).

During this phase of evolution the non-thermal radiation initially increases rapidly with decreasing radius ($\propto R^{-4} \sin^2 \chi$), and then decreases, to eventually become insignificant compared to L_{th} when the angular momentum loss due to the magnetic dipole radiation finally brings the magnetic and rotation axes into alignment (i.e. $\chi \longrightarrow 0$). In the regime when $L_{md} \gg L_{th}$, angular momentum loss by non-thermal radiation dominates that by equatorial mass loss, and Ozernoy & Usov (1973) show that χ changes according to the relation

$$\sin^2 \chi \sim 1 - \frac{R_{md}}{R} \cos^2 \chi_i \qquad (162)$$

where χ_i is the angle between the magnetic and rotation axes at the point when $R = R_{md}$, defined to be the radius at which $L_{md} = L_{th}$. Substituting this into (159) thus gives

$$L_{md} \sim 1{\cdot}4 \times 10^{52} \, \mathfrak{Z} \left(\frac{\Omega}{\Omega_{cr}}\right)^4 \left(\frac{R_g}{R}\right)^4 \left(1 - \frac{R_{md}}{R}\cos^2 \chi_i\right) \quad W \qquad (163)$$

The definition of R_{md} implies

$$\frac{3}{8}\,\frac{c^5}{G}\,\mathfrak{Z}\left(\frac{\Omega}{\Omega_{cr}}\right)^4\left(\frac{R_g}{R}\right)^4 \sin^2 \chi_i \sim \frac{4\pi G c}{\kappa}\,\mathcal{M}$$

$$\Rightarrow \quad R_{md} \sim \left(\frac{3\kappa G^2}{2\pi c^4}\right)^{1/4} \mathfrak{Z}^{1/4} \left(\frac{\Omega}{\Omega_{cr}}\right) \left(\sin \chi_i\right)^{1/2} \mathcal{M}^{3/4}$$

$$\text{i.e} \quad R_{md} \sim 1{\cdot}6 \times 10^{13} \, \mathfrak{Z}^{1/4} \left(\frac{\Omega}{\Omega_{cr}}\right) \left(\sin \chi_i\right)^{1/2} \left(\frac{\mathcal{M}}{10^6 M_\odot}\right)^{3/4} \quad m \qquad (164)$$

where again the opacity has been assumed to be caused by electron scattering in a medium with $X = 0{\cdot}73$. This expression should be compared with equation (6.12) of Ozernoy & Usov (1973), which is too small by the factor $10^{\frac{1}{4}}$.

The peak non-thermal luminosity may be shown to occur at radius $R_{L_{md,max}}$, given by

$$R_{L_{md,max}} = \frac{5}{4}\, R_{md} \cos^2 \chi_i \,. \qquad (165)$$

Substituting this into (163) thus gives

$$L_{md,\,max} = \left(5 \sin^2 \chi_i\right)^{-1} \left(\frac{5}{4}\cos^2 \chi_i\right)^{-4} L_{th} \qquad (166)$$

(Note that this relation is only true provided that (162) is valid, which in turn is based on the approximation $L_{md} \gg L_{th}$).

The qualitative evolution of the spinar depends on the angle χ_i ; and Ozernoy & Usov (1973) distinguish the following four cases:

<u>Case I</u> $R_{md} \leqslant R_{cr}$

In this case thermal radiation dominates the output of the spinar right up to the point at which stability is lost at $R \sim R_{cr}$. The angle χ_i lies in the range $0 \leqslant \chi_i \lesssim \chi_1$, where χ_1 is defined by the condition $R_{md} = R_{cr}$; i.e.

$$\sin \chi_1 \sim 6 \times 10^{-6} \zeta^{-1/2} \left(\frac{0 \cdot 1}{z}\right)^2 \left(\frac{\Omega_{cr}}{\Omega}\right)^2 \left(\frac{M}{10^6 M_\odot}\right)^{1/2} \tag{167}$$

This is a very small angle; for $z \sim \zeta \sim \frac{1}{4}$ and $M \sim 10^6 M_\odot$ it is only $\sim 0 \cdot 4$ arc seconds.

<u>Case II</u> $R_{cr} < R_{md} < R_{L_{md,max}}$

In this case the non-thermal luminosity increases with decreasing radius until, at $R \sim R_{md}$, the non-thermal and thermal powers are comparable. At smaller radii the magnetic and rotation axes begin to move together, and L_{md} is maintained at $L_{md} \lesssim L_{th}$. χ_i lies in the range $\chi_1 < \chi_i < \chi_2$, where

$$\chi_2 = \sin^{-1}\left(\frac{1}{\sqrt{5}}\right) \sim 26 \cdot 6^\circ \tag{168}$$

<u>Case III</u> $R_{cr} < R_{L_{md,max}} < R_{md}$

In this case L_{md} increases rapidly with decreasing radius until the point of maximum non-thermal output is reached. The peak luminosity is defined by (166), and χ_i lies in the range $\chi_2 < \chi_i < \chi_3$, where

$$(\sin \chi_3)^{1/2} . \cos^2 \chi_3 \sim 2 \times 10^{-3} \left(\frac{0 \cdot 1}{z}\right)\left(\frac{\Omega_{cr}}{\Omega}\right) \zeta^{-1/4} \left(\frac{M}{10^6 M_\odot}\right)^{1/4}$$

$$\Rightarrow \chi_3 \sim \frac{\pi}{2} - 0.045 \times \left(\frac{0.1}{z}\right)^{1/2} \left(\frac{\Omega_{cr}}{\Omega}\right)^{1/2} \zeta^{-1/8} \left(\frac{M}{10^6 M_\odot}\right)^{1/8} \qquad (169)$$

For $z \sim \zeta \sim \frac{1}{4}$ and $M \sim 10^6 M_\odot$, this is on the order of 88°.

<u>Case IV</u> $R_{L_{md,\,max}} \leqslant R_{cr}$

In this case the approaching of the magnetic and rotation axes only becomes of importance after the point of secular instability has been reached, so $L_{md} \gg L_{th}$ for $R_{cr} \lesssim R \lesssim R_{md}$.

The highest luminosities attained by the spinar during its quasi-static phase of evolution occur when the angle χ_i lies in the approximate range $26.6^\circ \lesssim \chi_i \lesssim 88^\circ$, indicating

$$1 \lesssim (L_{md,max}/L_{th}) \lesssim 4 \times 10^{10} \qquad (170)$$

This shows that spinars with large values of χ_i can in principle achieve very large non-thermal luminosities, although it should be noted that the upper limit on the RHS of (170) is a rapidly varying function of χ_i for $\chi \sim 90^\circ$. For example, if $\chi_i \sim 80^\circ$ then the peak non-thermal luminosity is of order $10^5 L_{th}$, and if $\chi_i \sim 75^\circ$ it is only about $4000 L_{th}$. The peak luminosities of $10^6 M_\odot$ spinars with $\chi_i \sim 80^\circ$ and 75° are thus $\sim 10^{42}$ W and $\sim 4 \times 10^{40}$ W respectively, which are comparable to the luminosities of the most powerful observed quasars. Typical Seyfert powers could be produced by lower-mass objects, or by similar spinars with smaller values of χ_i.

An important omission from the above discussion is the neglect of hydrogen burning. Following Fricke (1974) and Ozernoy & Usov (1977), the radius at which nuclear reactions first occur, R_{nuc}, may be estimated to be of order

$$R_{nuc} \sim 10^{12} \left(\frac{M}{10^6 M_\odot}\right)^{5/11} \quad m \qquad (171)$$

which, for $M \sim 10^6 M_\odot$, lies roughly mid-way between R_{qs} and R_{cr}. Comparison of R_{nuc} with R_{cr} shows that nuclear burning will occur before the onset of the relativistic instability for all masses less than M_{nuc}, given approximately by

$$M_{nuc} \sim 4 \times 10^8 \left(\frac{z}{0 \cdot 1}\right)^{11/6} \quad M_\odot \qquad (172)$$

An important parameter is the binding energy of the model. In the case of the 'hot' spinar, this is given (Fowler, 1971; Ozernoy & Usov, 1971) by

$$\frac{E_b}{Mc^2} \sim \frac{1}{4} K^2 \alpha^2 \left(\frac{R_g}{R}\right) - \zeta \left(\frac{R_g}{R}\right)^2 \qquad (173)$$

where $K^2\alpha^2$ is defined in terms of the parameter z introduced earlier by $z = \frac{1}{3} K^2 \alpha^2$, and ζ (here determining the first post-Newtonian correction) is given for a polytrope of index 3 by $\zeta = (9/8)^2$. Thus

$$\frac{E_b}{Mc^2} \sim \frac{3}{4} z \left(\frac{R_g}{R}\right) - \left(\frac{9}{8}\right)^2 \left(\frac{R_g}{R}\right)^2 \qquad (174)$$

Now (171) implies

$$\left(\frac{R_g}{R_{nuc}}\right) \sim 3 \times 10^{-3} \left(\frac{M}{10^6 M_\odot}\right)^{6/11} .$$

so the binding energy when nuclear burning first occurs is (for $M \sim 10^6 M_\odot$):

$$E_{b,nuc} \sim \left(2 \cdot 2 \times 10^{-3} z - 1 \cdot 1 \times 10^{-5}\right) M c^2 \qquad (175)$$

Thus, for typical values of z ($\approx 0{\cdot}1$, say), the binding energy when nuclear reactions first switch on is very much less than the amount of energy that they might be expected to release ($\sim 0{\cdot}7\ \%\ Mc^2$).

The effect of the onset of nuclear burning on the object during its quasi-static phase of evolution is impossible to determine without solving a detailed model: the possibilities (1) that nuclear reactions switch on non-explosively (e.g. Appenzeller & Tscharnuter, 1973), (2) that relaxational oscillations are produced (Fowler, 1966), and (3) that nuclear burning leads to total disruption of the body (e.g. Thorne, 1966, p.362; Von Hoerner & Saslaw, 1976), can not yet be distinguished theoretically. If nuclear burning does not totally destroy the body (either because nuclear reactions occur non-explosively or because energy is released during a series of relaxational oscillations), then eventual gravitational collapse seems to be unavoidable (e.g. Lynden-Bell, 1969; Rees, 1977). In this case the body either forms a massive black hole, or it evolves into a state not properly described by conventional physical theory. In the former of these cases, second and subsequent cycles of nuclear activity must be assumed to be due to accretion onto a black hole; in the latter case, it is not known whether or not the body's late evolution would leave a massive remnant. On the other hand, if nuclear burning does lead to total disruption of the body, then second and subsequent cycles of activity will be broadly similar to the first, and recurrent nuclear activity must then be explained by the recurrent formation and evolution of a low-mass spinar. Further implications of the present theory for theories of nuclear activity are considered in the next section.

VI-3 Discussion

It has been shown above that the first massive non-stellar object to form in the galactic nucleus (other than the cool gaseous disc) will probably have a mass in the range $10^4 \lesssim M/M_\odot \lesssim 10^6$, depending on the details of the model. If at the end of its evolution this body collapses either into a black hole or into a state not correctly described by conventional physical theory (but which still leaves a massive collapsed remnant), second and subsequent cycles of activity must be explained by accretion onto this object, and the problem of nuclear activity is reduced to determining the structure and evolution of the disc-remnant system. If on the other hand the body undergoes a disruptive explosion at the end of its evolution, either (on conventional theory) by the sudden release of large amounts of nuclear energy at an evolutionary state characterised by $E_b \lesssim 0{\cdot}7$ % Mc^2, or by the operation of 'new physics' in an as yet unexplored physical state, second and subsequent cycles of activity should be caused by similar processes to the first; in particular observed nuclear activity should be explicable in terms of the evolution of a low-mass body ($M \lesssim 10^6\ M_\odot$). These low masses imply that observed activity can not be explained by the evolution of non-magnetic supermassive stars (e.g. Hoyle & Fowler, 1963; Appenzeller & Tscharnuter, 1973; Fricke, 1973, 1974), since such objects radiate at only the Eddington limit. Even for $M \sim 10^6\ M_\odot$ this is only $\sim 10^{37}$ W; too small to explain typical Seyfert luminosities, quite apart from the most powerful quasars (with $L \sim 10^{41}$ W). For this reason, and also because the formation of a non-magnetic body in the nucleus does itself seem a rather unrealistic idealisation, it was assumed in the above discussion that the first massive object to form would be a rapidly rotating spinar. The non-thermal ('magnetic dipole') radiation from such an object is certainly capable of explaining the most luminous quasars, but present understanding of the detailed effects of the onset of nuclear burning does not yet allow the possibility of gravitational collapse to be distinguished theoretically from that of a disruptive nuclear explosion. Although the high-entropy spinar model of Ozernoy & Usov (1971, 1973) was described in some detail (to illustrate some of the compexities of developing a realistic evolutionary model), it does not seem that this particular spinar type can explain observed nuclear activity on conventional physical theory. This is because (on conventional theory) if such an object is the cause of recurrent nuclear activity, it is necessary that it be destroyed by a disruptive nuclear explosion when $R \sim R_{nuc}$. However, the binding energy of the 'hot' spinar at this stage is only of order $\sim 5 \times 10^{-4}\ Mc^2$. Thus, to explain typical Seyfert energies ($\approx 3 \times 10^{53}$ J) by the quasi-static evolution of this model requires $\sim 3000\ (10^6 M_{\odot}/M)$ cycles of activity, which indicates a mean interval between active periods of order 3×10^6 y. On such a model, signatures of past activity such as bursts of nuclear star formation would overlap. This is not supported by observation, so we conclude - if recurrence is accepted - that the high-entropy spinar models evolving entirely by conventional physical processes can be ruled out.

An important difference between high and low-entropy spinars, is that the disc types of a given mass and angular momentum are very much more tightly bound than nearly spherical configurations with the same values of M and J (e.g. Salpeter & Wagoner, 1971). Since the onset of nuclear burning in these models may also lead to disruption (e.g. Wagoner, 1971; Hara & Ikeuchi, 1976), it is possible that the binding energy at $R \sim R_{nuc}$ might be of order $E_{nuc} \sim 0{\cdot}7\ \%\ Mc^2$. A similar calculation to the one above, with $E_{b,nuc} \sim 0{\cdot}7\ \%\ Mc^2$, shows that $\sim 200\ (10^6 M_\odot/M)$ periods of activity are required to explain typical Seyfert energies, indicating a mean interval between active phases of order $\sim 5 \times 10^7$ y. Within the uncertainties of this estimate it is not possible to rule out the repetitive formation and evolution of 'cold' low-mass spinars as the explanation of observed nuclear activity, although it should be noted that the assumption $E_{b,nuc} \sim E_{nuc}$ has probably led to an underestimate of the required number of Seyfert phases. _If_ detailed models of low-entropy low-mass spinars do indeed show that nuclear burning can lead to total disruption of the object when its binding energy is of order $0{\cdot}7\ \%\ Mc^2$, it may be concluded that such objects _might_ explain most observed activity. On the other hand if explosive nuclear burning occurs when the binding energy is too small, the model will be subject to the same difficulties as the high-entropy spinar model described above. If nuclear burning occurs non-explosively, or when $E_b > E_{nuc}$, total disruption of the body is not possible, and a fraction of the mass must then collapse to form either a black hole or an object that _is_ able to explode disruptively, but by processes involving still undiscovered physical laws. In the former case, second and subsequent cycles of activity must be understood by processes involving accretion onto the compact object; in the latter case, the conclusion that observed quasar luminosities can be understood by the evolution of a low-mass object is still valid, but $E_b \sim E_{nuc}$ is no longer a limit on the amount of energy that might be extracted before the body's eventual destruction.

If the spinar does not undergo a disruptive explosion at the end of its evolution (either by processes involving 'conventional' or 'new' physics), the first object to form is likely to be a collapsed remnant with mass in the same range as that of the original spinar (i.e. M_h $10^4 - 10^6\ M_\odot$). If the remnant is a massive black hole (e.g. Lynden-Bell, 1969; Lynden-Bell & Rees, 1971; Pringle et al, 1973), it is possible that energy will be produced by infalling matter with an efficiency η of order 30 % (Thorne, 1974). However although it has been shown that the present theory is consistent with the black hole model (in that typical Seyfert luminosities might be understood in a natural way; e.g. equation (125)), it is not clear that the same model can successfully explain the most luminous quasars (e.g. Callahan, 1977). A difficulty of detail with such models is how plausibly to obtain the high central fluxes of matter required to produce the highest observed luminosities: For example, in the present theory rough estimates of the expected central flux may be obtained by use of (123), since it can be shown that g_{mag} (the couple due to magnetic torques) is of the same order as g_{turb} (the couple due to 'turbulent' friction). Specifically, following Lynden-Bell (1969, 1971) and Lynden-Bell & Pringle (1974), we have

$$g_{mag} \sim \pi \varpi^2 H \mathcal{E}_{mag}$$

and

$$g_{turb} \sim 3\pi \varpi \nu \sigma V_c(\varpi) .$$

where H is the disc width, $\mathcal{E}_{mag}$ is the magnetic energy density, and the circular velocity has been assumed to be Keplerian; i.e. $V_c(\varpi) = (GM_h/\varpi)^{\frac{1}{2}}$. Writing $\nu \sim \frac{1}{3}\bar{c}^2/K$, and substituting $K \sim V_c(\varpi)/\varpi$ and $\mathcal{E}_{turb} \sim \rho\bar{c}^2 \sim \frac{\sigma}{H}\bar{c}^2$ thus gives

$$\frac{\mathcal{G}_{mag}}{\mathcal{G}_{turb}} \sim \frac{\mathcal{E}_{mag}}{\mathcal{E}_{turb}}$$

which, for discs, whose magnetic and turbulent energies are in rough equipartition, is of order unity. Thus, the expected central flux is given approximately by (123); i.e.

$$F \sim F_{crit} \sim 0{\cdot}06 \left(\frac{\bar{c}}{20\,km\,s^{-1}}\right)^2 \left(\frac{V_{s,crit}}{1\,km\,s^{-1}}\right) \quad M_\odot\, y^{-1} \qquad (176)$$

which implies that turbulent velocities $\bar{c} \gtrsim 500$ km s^{-1} are required to fuel the most luminous quasars ($L \sim 10^{41}$ W) with inward fluxes $F \sim 60\ (0{\cdot}3/\eta)\ M_\odot\ y^{-1}$.

A more severe difficulty with the black hole hypothesis, especially when confronted with the observational and theoretical evidence presented here favouring recurrence, is the seemingly inescapable prediction that most massive galaxies should contain supermassive black holes in their nuclei (e.g. Ryan, 1972; Rees, 1978). In particular, a number of independent arguments lead to the conclusion that our own Galaxy should contain a massive black hole: First, if our Galaxy has ever been active in the sense of Seyfert or quasar activity, then the required mass of the black hole may be estimated roughly by equating its peak luminosity to the Eddington limit of a mass M_h (e.g. Bisnovatyi-Kogan & Blinnikov, 1977; Callahan, 1977). If our Galaxy has ever been through a typical Seyfert period producing a luminosity L_{Sy}, the required central mass is

$$\mathcal{M}_h \sim \frac{K}{4\pi G c}\, L_{Sy}$$

$$\text{i.e.} \quad \mathcal{M}_h \sim 7 \times 10^6 \left(\frac{L_{Sy}}{10^{38}\,W}\right) M_\odot \qquad (177)$$

Arguments based on the observed numbers of quasars (e.g. Lynden-Bell, 1969) lead to the expectation that our Galaxy might even have been a quasar at some time in the past; in this event (177) should be multiplied by the ratio $L_q/L_{Sy} \sim 50 - 100$, and it may be concluded (cf. Lynden-Bell & Rees, 1971) that our Galaxy ought to have a black hole of mass $10^7 - 10^9\ M_\odot$ in its nucleus. Secondly, studies of the tidal break-up of stars by massive black holes in galactic nuclei (e.g. Hills, 1975; Frank & Rees, 1976; Young, 1977) indicate that if the initial hole mass is $\gtrsim 10^6\ M_\odot$, in 10^{10} years in a normal galactic nucleus its mass could be expected to exceed $\sim 10^8\ M_\odot$. The growth timescale may be estimated (Hills, 1975) by assuming that the hole accretes matter at a rate equal to its Eddington limit; i.e.

$$\eta \dot{\mathcal{M}}_h c^2 \sim \frac{4\pi G c}{\kappa} \mathcal{M}_h$$

$$\Rightarrow \quad \mathcal{M}_h(t) \sim \mathcal{M}_h(0) \exp(t/t_E) \qquad (178)$$

where t_E is the Eddington limit growth timescale, defined by

$$t_E \sim \frac{\eta \kappa c}{4\pi G} \sim 1{\cdot}2 \times 10^8 \left(\frac{\eta}{0{\cdot}3}\right) \quad y \qquad (179)$$

and, as before, the opacity is assumed to be caused by electron scattering in a medium with $X = 0{\cdot}73$. Although the growth of low-mass ($\lesssim 10^6\ M_\odot$) black holes by tidal disruption of nearby stars is uncertain, and limited more by the supply of material than the Eddington limit (e.g. Frank, 1976; Frank & Rees, 1976), the arguments in Chapters IV and V show that the nuclear region of a typical spiral galaxy is expected to contain copious amounts of gas. Thus, especially during the early stages of galactic evolution ($t \lesssim 10^9$ y, say), Hills' approximation of mass-limited rather than supply-limited growth may be quite realistic. Even a $10^4\ M_\odot$ hole can grow to a mass $\sim 10^8\ M_\odot$ on a timescale of order 10^9 years, so again - if our Galaxy ever contained a $10^4\ M_\odot$ black hole at an early stage of its evolution - the present hole mass might now be expected to exceed $\sim 10^8\ M_\odot$. A third estimate of the expected black hole mass may be obtained by assuming that the black hole has accreted just sufficient material to explain the estimated energy of the Seyfert phenomenon; i.e.

$$\mathcal{M}_h \sim \frac{E_{sy}}{\eta c^2}$$

$$\Rightarrow \quad \mathcal{M}_h \sim 6 \times 10^6 \left(\frac{E_{sy}}{3 \times 10^{53}\ \mathrm{J}}\right)\left(\frac{0 \cdot 3}{\eta}\right) \mathcal{M}_\odot \qquad (180)$$

Finally, if the black hole accretes gas lost from stars, either continuously or intermittently (e.g. Shields & Wheeler, 1978), the accretion rate may be estimated to be of order $\alpha_* M_n \sim 10^{-2}\ M_\odot\ y^{-1}$ for our Galaxy. The expected hole mass after a galactic lifetime of order 10^{10} years is thus again expected to be $\sim 10^8\ M_\odot$. These arguments all lead to black hole masses which are larger than that inferred from observations (e.g. Shklovsky, 1975; Wollman et al, 1977; Dokuchaev & Ozernoy, 1977), and therefore (provided that our own Galaxy is indeed representative) make the black hole hypothesis difficult to maintain.

The principal conclusion to be drawn from this discussion is that the present theory leads naturally to the hypothesis that nuclear activity is powered by low-mass ($M \lesssim 10^6\ M_\odot$) disc-like spinars. Non-thermal radiation from such bodies can in principle easily explain the highest observed quasar luminosities, although if they are to avoid collapse into a black hole it is necessary that they explode disruptively at some point of their evolution. The black hole hypothesis is not supported by observations of our Galaxy, and it is therefore concluded that the end-point of evolution of such a spinar ($M \lesssim 10^6\ M_\odot$) is either a disruptive nuclear explosion (cf. Hoyle & Fowler, 1973) or disruption from a state more tightly bound than E_{nuc} by processes involving 'new' physics (e.g. Clube, 1978).

VII INTERPRETATION OF OUR GALAXY IN TERMS OF RECURRENCE

In this thesis it has been shown (Chapter II) that the available observational evidence supports recurrence more strongly than single activity. This interpretation can be understood most simply on a model in which stellar mass loss is the primary source of fuel for nuclear activity (Chapters III and IV), and a detailed theory based on this source of gas was described in Chapters V and VI. The theory led naturally to the conclusion that in the first instance evolution of the nuclear disc should give rise to the formation of a low-mass low-entropy spinar, but because the evolution of such objects still requires detailed theoretical investigation, the two possible evolutionary end-points (gravitational collapse into a black hole or explosive disruption) could not be distinguished theoretically. Observations of our own Galaxy did not support the hypothesis that a black hole was formed, and it was tentatively concluded that this possibility did not occur. The present theory has been based on parameters thought to be typical of our own Galaxy, and if this galaxy is indeed representative of the spiral class as a whole, it too has presumably been subject to recurrent nuclear outbursts. Because of this, and also because observations of our Galactic Centre place tight constraints on the black hole theory of nuclear activity, it is of great interest to determine whether or not detailed observations of our Galaxy lend support to the present ideas.

This chapter divides broadly into two parts. The first presents a brief review of the most important characteristics and features observed towards the Galactic Centre. It deals in turn with (A) infrared observations, (B) observations of the ionised gas,

(C) observations of the neutral gas, (D) observations of the star distribution, and finishes (E) with a short summary. The second part presents a qualitative interpretation of these observations in terms of recurrence, and it is concluded that the available evidence points towards our Galaxy having just finished a period of Seyfert activity.

VII-1 Review of observations of the Galactic Centre

VII-1A Infrared observations

The first infrared observations of the Galactic Centre were made by Becklin & Neugebauer (1968), who discovered a dominant extended near-infrared source coincident in position and extent with the radio source Sgr A. Its total projected area, elongated in the Galactic plane, was $\sim 30 \times 15$ pc (the full width half maximum (FWHM) was about half this), and it was found that the surface brightness decreased outwards from a central 'core' of FWHM ~ 1 pc roughly proportional to $R^{-0 \cdot 8}$ (All distances quoted here are based on an assumed distance to the Galactic Centre of 10 kpc). The spectrum of the source was compatible with that of a black body at $T \sim 900$ K, and since comparison with the central region of M 31 showed that the two galaxies had very similar near-infrared profiles, it was concluded that the emission from our own Galactic Centre was probably due to population II stars (at $T \sim 4000$ K) reddened by about 27 magnitudes of visual extinction. The radial dependence of surface brightness indicated that the star distribution had a central core of diameter ~ 1 pc, and that beyond this the density decreased outwards proportional to $R^{-1 \cdot 8}$.

Close to the centre of this dominant source (at a distance of $\sim 0 \cdot 5$ pc) they found a point-like source of $2 \cdot 2\ \mu$m emission with absolute magnitude ~ -11. This was interpreted as being probably a young massive star similar to α Ori which had recently formed and evolved close to the nucleus.

Apart from these features, the early $2 \cdot 2\ \mu$m maps showed, in a general background extending up to $\frac{1}{2}^{\circ}$ from the centre, that the nuclear star cluster extended outwards to a distance of at least 100 pc. Complicating and confusing the picture were also a number of other (extended) sources, and these were provisionally identified as parts of the dominant source seen through slightly less absorbing matter.

Since these early observations were made, new data obtained with higher sensitivity and resolution have confirmed the overall picture described above (e.g. Ito et al, 1977), although the structure of the central 'core' is now known to be extremely complicated. Within this region (of total diameter ~ 2 pc) most of the $2{\cdot}2\ \mu m$ emission comes from discrete generally unresolved sources with absolute magnitudes brighter than ~ -8 (Becklin & Neugebauer, 1975); the resolution of the most recent observations being $\sim 2{\cdot}3$ arc seconds, corresponding to a linear size of order $0{\cdot}1$ pc. The majority of these sources have now been identified as massive evolved stars, star clusters, planetary nebulae or compact H II regions, showing that our Galaxy has recently undergone a burst of nuclear star formation (Neugebauer et al, 1976; Becklin et al, 1978). One source however (IRS 16), although resolved, indicating a linear size $\sim 0{\cdot}1$ pc, is still unidentified. Its position (coincident within the observational errors with that of the compact non-thermal radio source detected using VLBI techniques; e.g. Lo et al, 1975) is thought to overlie the position of highest star density in the Galaxy, and the object is peculiar, if it is an H II region, in that it shows no localised $10\ \mu m$ emission from associated dust. This has led to the suggestion that perhaps it is a cluster of stars, but in this case, since it shows no CO absorption feature, the stars are either giants of spectral type earlier than M0, or dwarfs, in which case a very large number ($\sim 10^6$) are needed to produce the observed emission. In either case the indicated star density is very high (Becklin et al, 1978), and the cluster must be rapidly evolving on a short timescale ($T_{ref} \lesssim$ few x 10^5 y or $\lesssim 10^7$ y, depending respectively on whether the assumed model is giant or dwarf-dominated; cf. equation (1)). Due to it still being unidentified, the nature of IRS 16 is open to much speculation, and it has been suggested that it might be explained by a massive black hole surrounded by a cluster of M dwarfs.

Infrared observations at longer wavelengths (~ 100 μm) have been made by a number of authors (e.g. Hoffmann et al, 1971; Harvey et al, 1976; Gatley et al, 1977), and show a large extended source with dimensions in l x b of $\sim 630 \times 350$ pc, which can be understood as due to starlight re-radiated by dust. Close to the centre (i.e. within a radius of ~ 15 pc) the dust is thinly and fairly uniformly distributed, and does not decrease outwards as rapidly as the star density. The ratio by mass of dust to stars is $\sim 10^{-5}$ that in the solar neighbourhood, and the visual extinction across the central ~ 10 pc is only about 3 magnitudes (Gatley et al, 1977).

An important use of the far-infrared data has been to estimate the total luminosity coming from the Galactic Centre region, since any radiation emitted in the waveband $\sim 0{\cdot}1 - 1$ μm will be absorbed and re-radiated by dust. Comparison of the observed far-infrared luminosity with that estimated by other means (i.e. by summing the contributions inferred from sources such as stars and H II regions that are observed at near-infrared and radio wavelengths) thus enables an estimate to be made of whether or not an as yet undiscovered source is contributing to the power output. An important conclusion drawn by Gatley et al (1977) and Krugel & Tutukov (1978) is that no 'hidden' source of radiation is required to explain the data, although uncertainties in interpretation probably still allow the possible presence of a source producing $L \lesssim 10^7 L_\odot$ to be located somewhere within the very central region ($R \lesssim 1$ pc).

VII-1B Radio continuum and recombination line observations

Early observations (e.g. Mills, 1956; Smith et al, 1956) indicated the presence of at least two types of source in the direction of the Galactic Centre: (1) a discrete source observed at centimetre wavelengths with angular size $\sim \frac{1}{2}^{\circ}$ superposed on a fainter extended source; and (2) an extended source of size in l x b $\sim 5^{\circ} \times 2^{\circ}$ ($\sim$ 900 x 350 pc) that was seen predominantly at longer wavelengths ($\lambda \sim 3{\cdot}5$ m). This source, whose centroid was practically coincident with the discrete centimetre source, was seen to have two maxima separated by $\sim 2^{\circ}$ placed more or less symmetrically about the discrete source. Comparison of the flux at 3·5 m with that at 1·2 m showed that it decreased with decreasing wavelength, and that it was therefore non-thermal in origin. On the other hand, measurements of the discrete source at 9·4 cm and 3 cm showed that its spectrum was flat, consistent with it being thermal emission from an H II region situated either at the Galactic Centre or along the line of sight. This hypothesis also neatly explained the 'double' structure of the extended non-thermal source at $\lambda \sim 3{\cdot}5$ m: the intervening thermal plasma of the H II region absorbed the radiation coming from the non-thermal source and lowered its surface brightness just where its maximum should have been.

The extended non-thermal source, whose boundaries lie somewhere between the outer boundaries of the spheroidal bulge component of our Galaxy constructed by Sanders & Lowinger (1972) and the observed extent of the far-infrared radiation, is now thought to be diffuse synchrotron emission similar to that observed throughout the Galaxy; its enhancement towards the Galactic Centre is due either to an enhanced magnetic field or to an enhanced density of relativistic electrons in that region.

Recent observations have considerably sharpened our picture of the source seen at centimetre wavelengths, and it is now known that the dominant contributor to the centimetre flux from a region of dimensions $l \times b \sim 250 \times 90$ pc, centred on the nucleus, is thermal free-free emission from an extended H II region with mean electron density $n_e \sim 15\ cm^{-3}$, temperature $T \sim 10^4$ K and total ionised mass $M_{H\ II} \sim 10^6\ M_\odot$ (Mezger, 1974; Table I). Observations of recombination lines from the same region confirm the thermal nature of the source, and show the presence of random velocities up to of order 100 km s^{-1} (e.g. Mezger et al, 1974). Embedded within this extended H II region are a number of discrete radio sources, most of which are thought to be compact H II regions of somewhat higher density than the extended source (i.e. $n_e \sim 100 - 200\ cm^{-3}$).

Close to the centre of the extended H II region is the discrete radio source designated Sgr A. At centimetre wavelengths it shows a typical non-thermal spectrum which flattens with increasing wavelength and eventually turns over at $\lambda \sim 1{\cdot}5$ m (Mezger et al, 1974). This kind of behaviour can be understood on a model in which a non-thermal source is observed through a thermal plasma, and recent recombination line observations (Pauls et al, 1974) from a region of size $l \times b \sim 11 \times 13$ pc surrounding the nucleus have lent support to such a picture. At $\lambda \sim 6$ cm the thermal plasma accounts for $\sim 41\ \%$ of the flux density from the source, and Pauls et al have derived an electron density $n_e \sim 200\ cm^{-3}$ and total ionised mass $\gtrsim 10^4\ M_\odot$ for the region. Within Sgr A lie the two other principal contributors to the 6 cm flux: Sgr A East (44 %) and Sgr A West (15 %). The first to report that they had resolved Sgr A into these two components were Downes & Martin (1971), and it was shown that the eastern component was non-thermal (now believed to be a supernova remnant), whereas the western component, overlying the near-infrared 'core', was a still more compact thermal source.

Sgr A West has a radius a little over 1 pc, and is a dense H II region surrounding the region of highest star density in the Galaxy (Downes & Martin, 1971; Ekers & Lynden-Bell, 1971). Its electron density is estimated to be $\sim 1400\ \mathrm{cm}^{-3}$ (Mezger, 1974), and recombination line studies indicate the presence of very large gas motions with $\Delta V \sim 200\ \mathrm{km\ s}^{-1}$ (Pauls et al, 1974, 1976), the origins of which are still unknown. Similar line widths have also been observed in the infrared by means of the $12 \cdot 8\ \mu\mathrm{m}$ [Ne II] line (Wollman et al, 1976, 1977).

Within Sgr A West is radio fine structure on a scale of order $\lesssim 0 \cdot 5$ pc, tentatively identified by Balick & Sanders (1974) as thermal emission from dense compact H II regions embedded within the Sgr A West source itself; the density in these regions being estimated to be $n_e \gtrsim 10^4\ \mathrm{cm}^{-3}$, and their contribution to the total emission from Sgr A West $\sim 10 - 20$ % (Balick & Sanders, 1974).

Using still higher resolution, and wavelengths of 11 cm and $3 \cdot 7$ cm, Balick & Brown (1974) detected a still unresolved source close to the centroid of the infrared 'core' at the centre of the Galaxy. Its small angular diameter (corresponding to $\lesssim 5 \times 10^{-3}$ pc) and very high brightness temperature ($T_b \gtrsim 10^7$ K) suggested a compact non-thermal source, although because its observed radio luminosity ($\sim \frac{1}{4} L_{\odot}$) was very much less than that of similar compact sources detected in the nuclei of other galaxies (e.g. Kellermann et al, 1976), its detailed relation to possible nuclear activity remains unclear. The source has since been observed a number of times (e.g. Lo et al, 1975; Davies et al, 1976; Kellermann et al, 1977), and it is now clear that a part of it is very small indeed. Kellermann et al (1977) show that about 25 % of the emission comes from a region only about 10 AU across ($\lesssim 5 \times 10^{-5}$ pc), and if Davies et al (1976) are correct in arguing that its apparent diameter is due largely to the effects of interstellar scattering, even this size is an upper limit. Various suggestions have been made as to the object's nature, ranging from the idea that is is simply a supernova, pulsar or some other kind of galactic star (see Davies et al, 1976), to the suggestion that it is related in some way to the kind of activity observed in other galaxies (e.g. Kellermann et al, 1977), or that it is a massive black hole surrounded by a dense stellar cusp (Thorne & Braginsky, 1976).

VII-1C Observations of the neutral gas

The neutral material towards the Galactic Centre consists of gas which is both in molecular and in atomic form. The most widely distributed is atomic hydrogen, observed from the whole region $0 \lesssim \varpi \lesssim 4$ kpc, and detected both in emission and absorption against the continuum radio sources close to the centre. The molecular gas is mostly confined to dense clouds within the central ~ 300 pc. Molecular hydrogen has not been observed directly, but its presence is inferred from observations of other molecules, in particular carbon monoxide (CO) and ammonia (NH_3), seen in emission; and hydroxyl (OH) and formaldehyde (H_2CO), seen in absorption. Observations of molecules in emission and absorption are complementary: In the former case, although the cloud structure is quite easily defined, it is difficult to assign correct spatial positions to the observed features; in the latter case (OH and H_2CO), although there is not usually any ambiguity as to whether the observed cloud lies in front of or behind the Galactic Centre, a correct interpretation of the observed line strengths must take proper account of the spatial variation of the underlying continuum source.

One of the first attempts to construct a kinematic model explaining the molecular features was that of Scoville (1972), who suggested that the observations could be understood in terms of a model in which the clouds were confined to a series of expanding and rotating rings. The alternative hypothesis, that the clouds were simply unrelated features moving in independent elliptical orbits, was ruled out on the ground that many appeared to show 'bridges' between one another, indicating that they formed part of the same physical structure. Even so, since uniform rings could not explain the data, some degree of 'clumpiness' was required. Scoville placed the clouds at distances of order 250 pc from the nucleus, and estimated their rotation and expansion velocities to be of order 50 km s^{-1} and 100 km s^{-1} respectively.

A serious difficulty with any expansion model is the so-called '+ 40 km s^{-1} feature'; a dense molecular cloud seen in absorption against Sgr A moving away from the sun at a velocity of $\sim$ 40 km s^{-1}. Recent observations (e.g. Sandquist, 1974; Rieke et al, 1978) show that this cloud extends weakly in front of all the Sgr A components, so the explanation preferred by Oort (1974a, 1977), that the cloud is due to expulsion from Sgr A West in a direction away from the sun in such a way that it just happens to lie in front of Sgr A East (presumed to lie on the far side of the Galactic Centre), is no longer tenable. Sandquist (1974) has suggested that the cloud is physically associated with the nucleus and that it consists of a rapidly rotating cloud of gas and dust with radius R $\sim$ 10 pc and a large ($\sim$ 50 km s^{-1}) inward component of velocity. Clube (e.g. 1978) has also identified this feature as physically associated with the nucleus, but instead ascribes its large positive velocity to a net outward motion of the local standard of rest, rather than to infall.

Almost simultaneously with the appearance of Scoville's work, Kaifu et al (1972) published a different interpretation of the molecular line data, and proposed two types of ring: one expanding, with parameters similar to those suggested by Scoville; and the other, to explain the 40 km s^{-1} feature as on the near side of the Galactic Centre, contracting. This ring was assumed to have an inward component of velocity of order 40 km s^{-1}, and a radius $\varpi \sim$ 140 pc. They estimated the total mass of molecular gas in the expanding ring to be very large ($\sim 10^8 - 10^9$ $M_\odot$), but due mainly to uncertainties in estimating the filling factor and the $n(CO)/n(H_2)$ ratio (e.g. Oort, 1977), this estimate has since been revised downwards to $\sim 10^7$ $M_\odot$. Kaifu et al (1974) have further developed this interpretation, and suggest that the appearance of expanding and contracting rings can be understood physically in terms of a shock-wave model in which the inward moving shock is caused by gas displaced by the expanding shock returning to its original position. Work by Mezger et al (1974) has however cast doubt on the reality of the contracting ring, and if Sandquist's (or even Clube's) explanation of the 40 km s^{-1} feature is accepted, then the inner ring described by Kaifu et al (1974) is probably unnecessary to explain the observations. An interesting trend in the data, noted by Kaifu et al (1974), is that the expanding molecular ring seems to be tilted with respect to the Galactic plane; and comparison with other features composed mainly of neutral gas led them to conclude that the nuclear gas was generally confined to a plane inclined at an angle of $\sim 6^\circ$ with respect to the plane $b = 0^\circ$. This trend has since been extensively confirmed by H I observations Cohen & Davies, 1976; Oort, 1977).

In contrast to the systematic outward motions revealed by the molecular data, early observations of atomic hydrogen in the central region ($\varpi \lesssim 800$ pc) revealed the presence of a rapidly rotating nuclear disc feature, showing no clear evidence for outward expansion (Rougoor & Oort, 1960; Rougoor, 1964). Circular velocities determined from this feature have traditionally been the means of probing the gravitational field in this region (e.g. Oort, 1971, 1977; Sanders & Wrixon, 1973) and there is considerable reluctance to accept an outward component of velocity for this feature, although radial motions as high as ~ 100 km s^{-1} can not be ruled out by the data (Cohen & Davies, 1976). However, although observations by Sanders & Wrixon (1973) could be understood in terms of an axisymmetric model rotating in centrifugal balance in the region $300 \lesssim \varpi \lesssim 800$ pc, within $\varpi \sim 300$ pc an expanding ring-like structure had to be invoked, and this was identified as coinciding partly with the expanding ring inferred from the molecular data. Sanders et al (1977) have further developed non-axisymmetric models of the nuclear disc, and have found that a spiral density enhancement is necessary to account for some of the observed features. Recently Cohen & Davies (1978) have suggested a similar model, with the important difference that the spiral 'arm' (containing perhaps a quarter of the mass) has a strong radial component of velocity similar to that revealed by the high resolution (telescope beam width ~ 6 pc) CO data discussed by Scoville et al (1974). These observations, in addition to detecting distributed CO emission from the nuclear disc feature, also provided further evidence for non-circular motions throughout the nuclear region, and Scoville et al (1974) argued in favour of a unified picture in which most of the nuclear gas was confined to an expanding two-armed spiral pattern. The total mass of molecular gas within the central ~ 600 pc was estimated to be in the range $10^7 - 10^8$ $M_\odot$, which, although an order of magnitude higher than the H I mass within the same region ($\sim 4 \times 10^6$ $M_\odot$; Sanders & Wrixon, 1973; Cohen & Davies, 1976; Oort, 1977), is in rough agreement with other estimates (e.g. Kaifu et al, 1975; Oort, 1977).

The gas distribution at distances $\varpi \gtrsim 500$ pc is dominated by that of atomic hydrogen (e.g. Oort, 1977), and presents an extremely complicated picture (e.g. Mezger, 1974; Cohen, 1975; Cohen & Davies, 1976; Oort, 1977, 1978). In this region ($\varpi \lesssim 4$ kpc, say) most of the mass is concentrated into a small number of expanding arm-like features (such as the '3-kpc' arm and the '135 km s^{-1}' arm) which have outward components of velocity in the range $\sim 50 - 175$ km s^{-1} and a total H I mass of order $5 \times 10^7 M_\odot$. The bulk of this mass is concentrated in the two arms referred to above; the remainder (making up a total H I mass of order $\sim 10^7 M_\odot$) comprises less massive features that are distributed by mass in and out of the plane in the approximate ratio 4 to 1 (e.g. Cohen & Davies, 1976; Oort, 1977). This gas too partakes in the general l-b tilt of the nuclear gas distribution (Cohen & Davies, 1976), and its total mass is estimated to be of order twice its H I mass (Oort, 1977). The total kinetic energy associated with the outward motion, assuming a mean outward component of velocity of order 100 km s^{-1} and a total mass $\sim 10^8 M_\odot$, is on the order of $\sim 10^{48}$ J (cf. Cohen & Davies, 1976; Table III); an energy that is not particularly large by Seyfert standards, but large enough to warrant an explanation in terms of a large-scale influence (such as nuclear activity or action of a bar) that could significantly affect the whole of the central region (e.g. Van der Kruit, 1971; Peters, 1975; Cohen & Davies, 1976, 1978; Cohen & Few, 1976; Oort, 1977; Clube, 1978).

VII-1D The star distribution

An important parameter that must be introduced at an early stage into any detailed model of the central regions of the Galaxy ($r \lesssim 1$ kpc, say) is the assumed star distribution $\rho_*(r)$. Optical observations of stars close to the centre are not possible, due to the large visual extinction in that direction (~ 27 magnitudes; Becklin & Neugebauer, 1968) consequent upon the sun's present position close to the heavily obscured Galactic plane, and most current estimates of the mass distribution have to rely on indirect arguments. In particular, the most commonly used (e.g. Becklin & Neugebauer, 1968; Mezger, 1974; Oort, 1971, 1977; Sanders & Lowinger, 1972) are (1) observations of the near-infrared surface brightness, assuming that this accurately reflects the projected star distribution, and (2) observations of the H I nuclear disc feature, assuming that this gas is rotating in centrifugal equilibrium. Although the assumptions underlying these arguments have still to be proved (and might prove to be false; e.g. Andriesse & de Vries, 1976; Cohen & Davies, 1976), because the models that have been constructed on this basis give results similar to that observed in the bulge of M 31, the spheroidal model constructed by Sanders & Lowinger (1972) (and which often forms the basis for discussion; e.g. Mezger, 1974; Oort, 1977) will be described here. It will be recalled that the assumed star density (14) was based on their work.

Sanders & Lowinger (1972) constructed a spheroidal model based on the elliptical contours of the near-infrared surface brightness distribution, and converted luminosities to masses by making the reasonable assumption that the stars in the bulge of M 31 were similar to those in the central region of our Galaxy. Specifically, adopting a mean mass-to-total-luminosity ratio of $M/L \sim 3$, they obtained

$$\rho_* \sim 7.6 \times 10^5 \, \alpha^{-1.8} \quad M_\odot \, pc^{-3} \qquad (181)$$

where α (measured in parsec) is an ellipsoidal coordinate related to ϖ and z (distances measured in and perpendicular to the plane respectively) by

$$\alpha^2 = \varpi^2 + z^2(1-e^2)^{-1} \qquad (182)$$

Here e is the eccentricity of the spheroidal distribution, taken to be $e = 0{\cdot}91$ to give an axial ratio $c/a \sim 0{\cdot}4$. The distribution was assumed to extend outwards as far as $R_{max} \sim 800$ pc, and inwards to $R_{min} \sim 0{\cdot}5$ pc. Within this radius, by analogy with the observed variation of the near-infrared surface brightness, the star density was assumed to be nearly constant. The total stellar mass of the spheroid was $\sim 1{\cdot}1 \times 10^{10}\ M_\odot$ (cf. $M_n \sim 10^{10}\ M_\odot$ used throughout this thesis), and the mass within a central region of radius ~ 1 pc was $\sim 3{\cdot}6 \times 10^6\ M_\odot$. The mass within the central 'core' of radius $R_{min} \sim 0{\cdot}5$ pc was $(1 - 2) \times 10^6\ M_\odot$, depending on the assumed eccentricity of the distribution close to the centre.

In recent years it has become possible to make an independent estimate of the mass within the central region using observations of the broad $12{\cdot}8\ \mu m$ [Ne II] emission from ionised gas close to the centre. By assuming that the line width is caused by large-scale motions of roughly the same magnitude as the circular velocity within radius ϖ, Wollman et al (1976, 1977) have estimated the total mass within $\varpi \sim 1$ pc to be of order $4 \times 10^6\ M_\odot$. This agrees well with the Sanders-Lowinger estimate of the same quantity, but such good agreement between the two independent estimates is probably fortuitous in view of many likely sources of error. However the result does indicate (cf. Section VI-3) that our Galaxy probably does not contain a very massive black hole. Oort (1977, 1978) has interpreted the [Ne II] data as evidence that the total mass within $R \sim 0{\cdot}4$ pc is of order $5 \times 10^6\ M_\odot$, exceeding the stellar mass within the same radius by a factor of 2 – 3; but even if this (extreme) interpretation of the neon data is accepted, the conclusion that this provides the first evidence for the possible existence of a massive black hole ($M_h \sim 3 \times 10^6\ M_\odot$) is not strongly indicated. This is because the discrepancy (of order $3 \times 10^6\ M_\odot$) depends not only on the 'neon-mass' (which is probably uncertain by at least a factor of two) but also on the assumed stellar mass, which was calculated by use of a mass-luminosity ratio determined by comparison with the bulge of M 31. In M 31 there is some evidence that its 'core' stars might differ systematically from those in its bulge component (e.g. Light et al, 1974; Morton & Elmegreen, 1976), so the assumption that M/L in the bulge is the same as that in the near-infrared core of our Galaxy is questionable. Even if the similarity is accepted however, recent determinations of the M/L ratio using population models indicate $M/L_V \sim 15$ (e.g. Oort, 1977) which is in good agreement with the independent estimates $M/L \sim 8 - 9$ obtained from velocity dispersion measurements (e.g. Light et al, 1974; Morton & Elmegreen, 1976). Adopting these more recent determinations of M/L increases the Sanders-Lowinger masses by a factor of 2 – 3; again consistent with the [Ne II] data without the need for a massive black hole.

In view of the various uncertainties, probably all that can be concluded reliably is that the mass of a hypothetical black hole certainly does not dominate the stellar mass within $\Theta \sim 1$ pc, and that the observations are consistent with $M_h \lesssim 10^6 M_\odot$, with zero remaining a real possibility.

VII-1E Summary

Close to the Galactic Centre is a still unresolved compact non-thermal radio source with linear dimensions $\lesssim$ 10 AU and radio power $\sim 10^{26}$ W. Its nature remains uncertain: It is possible that it is generically related to the kind of nuclear activity observed in other galaxies; alternatively it could simply be a young pulsar, supernova remnant or some other kind of galactic star. The position of the radio source coincides within the observational errors with the peak in surface brightness of IRS 16, which is a small (diameter $\sim$ 0·1 pc), barely resolved near-infrared source that could be either a compact H II region containing an unusually small amount of dust, or a dense compact star cluster. If the latter, its 2·2 μm power indicates a luminosity $L \lesssim 5 \times 10^4 L_\odot$ (Becklin et al, 1978; Rieke et al, 1978). Although the positional coincidence with the compact radio source is striking, and has been interpreted as evidence for a stellar cusp surrounding a massive black hole (Thorne & Braginsky, 1976), the latest near-infrared position measurements, showing that the peak of IRS 16 may be displaced from the radio source by as much as 0·05 pc (Becklin et al, 1978), do not support this interpretation. [Ne II] velocity dispersion measurements also do not support the hypothesis that our Galaxy contains a massive black hole in its nucleus, and a rough upper limit to the maximum allowable hole mass is $\sim 10^6 M_\odot$.

Surrounding these sources, and contained within the dense nuclear H II region Sgr A West, are a number of other near-infrared sources variously identified as compact H II regions or late-type supergiants. This shows that our Galactic Nucleus has recently experienced a burst of massive star formation (Neugebauer et al, 1976;

Rieke et al, 1978). The total luminosity from the Sgr A West region (radius ~ 1 pc) is estimated to be $\sim 3 \times 10^6\ L_\odot$, although the presence of a hypothetical source of optical or ultraviolet emission with total luminosity $\lesssim 10^7\ L_\odot$ can not yet be ruled out (Gatley et al, 1977). Sgr A West overlies the region of highest star density as inferred from near-infrared observations, and is embedded within a more extended lower-density complex of ionised gas designated Sgr A. This radio source contains gas with mass motions of order 200 km s^{-1}, molecular clouds, and a supernova remnant (Sgr A East); the mean electron density is $\sim 200\ cm^{-3}$, and the diameter ~ 10 pc.

Surrounding Sgr A is a much more extensive H II region (Mezger, 1974) containing a total ionised mass $M_{H\ II} \sim 10^6\ M_\odot$ at a mean electron density $\sim 15\ cm^{-3}$. The radius of this region is ~ 125 pc, and recombination line studies (e.g. Mezger et al, 1974) indicate large-scale mass-motions with velocities of order 100 km s^{-1}. In addition to containing a number of higher-density compact H II regions (such as Sgr A), this extended region of ionised gas also contains several discrete molecular clouds (Sanders & Wrixon, 1974), and, further from the centre, an increasing proportion of neutral gas. Within a radius of order 500 pc most of the neutral gas is in molecular form and distributed as an expanding ring-like or spiral pattern. A small fraction (~ 10 %) is in atomic form, and may be distributed as a rapidly rotating nuclear disc. The outward velocity of expansion of the molecular clouds is of order 100 km s^{-1}, close to the r.m.s. velocity dispersion indicated by recombination line observations from the same region.

At large distances from the centre ($\varpi \gtrsim 1$ kpc) the gas distribution is dominated by the outward expansion of arm-like features or ring-segments, and the whole of the nuclear gas distribution is observed to be inclined at a small angle $(\sim 4^{\circ} - 8^{\circ})$ with respect to the Galactic plane. The total mass in the outwardly moving 'arms' is estimated to be $\sim 10^{8}\ M_{\odot}$; i.e. about ten times larger than the mass contained in the inner expanding features. The total kinetic energy of outward motion is of order 10^{48} J.

VII-2 Discussion

Perhaps the strongest observational evidence in support of a theory of nuclear activity along the general lines of the one described in this thesis is the observation that our own Galaxy actually does contain a gaseous mass of order 10^8 $M_\odot$ distributed in atomic or molecular form within the central region $\varpi \lesssim 1$ kpc. If this gas was distributed uniformly as a quiescent nuclear disc, cooling in the disc would be so efficient that stellar mass loss from bulge-component stars would necessarily flow inwards, and mass-growth of the kind outlined in Section VI-1A would occur. At some point the gaseous disc must become gravitationally unstable, and the arguments given in Chapter VI then show that a period of nuclear activity is likely to occur, following the rapid inward viscous flow caused by turbulence generated by a burst of massive star formation. The timescale for formation of an active nucleus (in the first instance, the formation of a low-mass spinar) by this process may be estimated to be on the order of 10^6 y (cf. equations (24) and (144)). Since this is very much less than the timescale for shock-induced star formation to affect the whole of the nuclear disc (cf. Section II-1B), galaxies showing nuclear star formation combined with a relatively quiescent nuclear disc may be interpreted as being in a 'pre-active' state. Our Galaxy, although having large quantities of cool gas distributed throughout its nuclear region, does <u>not</u> appear to contain a dense quiescent nuclear disc, and this observation is consistent with the view that the initial burst of star formation occurred some time ago and has already led to a period of nuclear activity whose results we are now witnessing. The outwardly moving massive gas clouds, signs of nuclear star formation and motions out of the plane all lend support to this view, and it has been emphasised many times (e.g. Van der Kruit, 1971; Oort, 1977; Clube, 1978) that the wide-ranging variety of structures seen towards the Galactic Centre are extremely difficult, though perhaps not impossible (Peters, 1975), to explain without invoking recent nuclear activity. It is thus concluded that the general characteristics of our Galactic Centre can indeed be understood in terms of the present theory, and this implies that our Galaxy has recently finished a period of Seyfert activity. It is interesting to note that our Galaxy also shows signs of past activity as discussed in Chapter II, and this lends support to the overall self-consistency of the argument.

Although it is possible to argue that observations of our Galaxy are consistent with recurrence and the view that a period of nuclear activity has recently ended, a number of astronomers have in recent years taken the stance that it is quite possible that the same data could be understood on an entirely opposite view; namely, one interpreting the outwardly moving features as evidence for a central bar-like structure in the nuclear region (e.g. Simonson & Mader, 1973; Peters, 1975; Matsuda & Nelson, 1977; Cohen & Davies, 1976, 1978). Oort (1977) has discussed this question in some detail, and concludes that although the expanding arms distant from the centre might in principle be explicable in terms of a bar-like structure on a kiloparsec scale, the variety of expanding features closer to the centre, in particular those lying out of the plane, do not find a natural explanation on such a model. Three points should be emphasised. Firstly, no theoretical model (either assuming nuclear activity or a central bar) has yet been developed to the point where observations can reliably distinguish the hypotheses. Secondly, apart from the outward gas motions, no compelling independent evidence for a bar has been presented; and thirdly, even if a bar should eventually be shown necessary to explain some of the observed features, this in no way rules out the hypothesis of nuclear activity to explain others. Some Seyferts are observed to be barred spirals, so the hypotheses should perhaps be combined rather than confronted (e.g. Cohen & Few, 1976). If the Galaxy is a barred spiral it is not clear how a detailed theory of the disc's evolution would proceed, although, since barred types are not significantly more likely to be Seyferts than ordinary spirals (Adams, 1977), it is possible that close to the centre the differences might not be large. In any case, the theory developed here should still apply to ordinary spiral galaxies, and observational support for the general picture might then be obtained by comparison with other apparently post-active galaxies, such as NGC 253 (e.g. Ulrich, 1978).

A recent analysis by Clube (1978) has, however, raised the issue of deciding the origin of the expansion to a very high level of importance. In this work he argues that a large number of observations are consistent with the view that the whole Galaxy (or, more exactly, the whole solar neighbourhood) partakes in the general outward motion, and suggests that this can only be understood by invoking 'new physics' in nuclear activity. It should be noted that the present theory in no way rules this out (Section VI-3), although it has not yet been shown conclusively that more conventional explanations for the sun's apparent outward motion (e.g. action of a bar, or perhaps a tidal interaction with the Large Magellanic Cloud) can be ruled out.

In summary, the present chapter has shown that a wide variety of observations of our Galaxy can be easily incorporated into the present theory of nuclear activity. The available evidence, in addition to confirming the expectation that large quantities of gas should occur in the nuclear region, favours the interpretation that the Galaxy has recently finished a period of nuclear activity.

VIII CONCLUSIONS

The primary aim of the present thesis - reflected in its title - has been to try and explain observed nuclear activity as due to the inevitable evolution of ordinary galactic nuclei. This approach was governed primarily by the observational considerations discussed in Chapters I - III, where it was shown not only that a wide variety of evidence points towards the hypothesis that nuclear activity does occur in seemingly quite normal galaxies, but also that periods of Seyfert activity occur repeatedly during the evolutionary lifetime of massive spiral galaxies like our own. The idea that nuclear activity is a recurrent phenomenon is of course not new; it has been mentioned in the literature many times in the past (e.g. Burbidge & Hoyle, 1963; Ozernoy, 1974). Chapter II of this thesis however represents the first systematic analysis of the evidence in favour of recurrence, and it is also the first time that the potential importance of recurrence as a discriminant between different theories has been emphasised. For example, recurrence - if accepted on observational grounds - places theories requiring special or extreme 'initial conditions' for their success in severe difficulty (e.g. those theories which relate nuclear activity to processes occurring during the early stages of galactic evolution); and enables theories which rely solely on stellar dynamical processes, implying a single extended period of activity, to be ruled out (e.g. Spitzer & Saslaw, 1966; Hills, 1975). Adoption of recurrence almost certainly limits potential explanations of nuclear activity to one of two models; namely (1) activity is caused by intermittent accretion of gas onto a black hole, or (2) by the repeated formation and evolution of some other kind of massive object (e.g. a spinar).

In either case it seems likely that a disc of the type discussed in Chapters V and VI will play an essential role in the galaxy's evolution.

The principal conclusion to be drawn from the thesis is that a plausible sequence of events <u>has</u> been described which shows how a typical spiral galaxy such as our own must necessarily generate recurrently active nuclei. It is not yet possible to <u>prove</u> the argument, as many important aspects of the problem still remain only partially understood; however it is important to emphasise the overall coherence and self-consistency of the proposed evolution. If nuclear activity is indeed unable to remove all the gas from the central regions of galaxies (as is indicated by observations of Seyfert galaxies and post-active galaxies such as our own), once the activity associated with a particular Seyfert event has finished, it is extremely difficult to see how further mass-growth and evolution of the nuclear disc towards a second cycle of activity can be avoided (Chapter V). One cycle of the disc's evolution may be divided broadly into three phases: (1) an extended period of slow mass-growth with negligible viscous evolution; (2) the onset of gravitational instability and massive star formation; and (3) the generation of a large viscous flux, leading to a period of nuclear activity caused either by accretion onto a massive black hole or by the evolution of a low-mass ($M \lesssim 10^6\ M_\odot$) spinar.

An important conclusion to be drawn from the discussion in Chapter IV is that contrary to the assumption of many astronomers (e.g. Larson, 1974b; Faber & Gallagher, 1976; Gisler, 1976), thermally steady supernova-heated galactic winds need not always be out-flowing. The presence of a thermally unsteady flow during an early phase of galactic evolution ($t \lesssim 10^9$ y; Larson, 1974b) will lead to the formation of a dense nuclear disc in which dissipation is so effective that even hot infalling material may be absorbed without leading to a significant rise in temperature. Such a disc will behave as a sink with respect to any hot (supernova-heated) gas, and at later epochs, provided that the disc mass exceeds a certain critical value, the flow will continue to be directed inwards. Stellar mass loss from the inner bulge components of galaxies is thus not removed from the system, and gas must therefore accumulate near the galactic centre until it eventually either condenses into stars or contributes towards the formation of an active nucleus. The idea that stellar mass loss might fuel nuclear activity has been mentioned many times in the literature (e.g. Shklovsky, 1971; Spitzer, 1971; Ozernoy, 1974; Van den Bergh, 1975), but this is the first time that the problems raised by the calculations of Mathews & Baker (1971) have been faced and adequately dealt with (Section IV-2). It may also be noted that although the present investigation has been concerned almost entirely with spiral galaxies and their relation to the Seyfert phenomenon (using our own Galaxy as an example), the extension of the argument to ellipticals is quite straightforward. If ellipticals also contain massive molecular discs, the absence of atomic hydrogen (e.g. Faber & Gallagher, 1976) need no longer be interpreted as support for the galactic wind hypothesis.

The detailed discussion of the disc's structure given in Chapter V, although only an order of magnitude investigation, made possible the proof of two crucial steps in the argument. First, even quite low-mass discs are sufficiently dissipative to act as sinks even for hot infalling material; and secondly, equilibrium temperatures of discs with realistic masses ($\gtrsim 10^7\ M_\odot$, for Case A rotation laws) were very low, leading to the conclusion that most of the gas -

especially that close to the centre - would be mainly in molecular form. The first condition was necessary in order to ensure that stellar mass loss would flow inwards, and the second ensured that random gas motions were sufficiently small that the kinematic viscosity ($\nu \propto \bar{c}^2$) was small enough for viscous evolution during this slow growth phase to be neglected.

The disc was shown to grow slowly in mass until the onset of gravitational instability. In principle, this might first occur anywhere throughout the disc, but for the cases discussed in Chapters V and VI the disc was found to fragment first either at or close to the origin. The first bodies to form were found to have stellar masses, and it was assumed that ordinary stars, with the usual range of masses, would be formed. The possibility that star formation in the disc would always lead to low-mass stars, although conceivable, was rejected as unlikely. Formation of even a few massive stars will lead to further massive star formation (Elmegreen & Lada, 1977), and it was assumed that this process would continue until - after a relatively short period - the entire central region of the disc was converted into a highly disturbed state characteristic of other star-forming regions in galaxies. The viscosity would by then have increased by a very large factor, and large inward fluxes would be produced.

Chapter VI dealt in a more qualitative way with the likely late evolution of the disc, and considered various implications of the theory for detailed models of active nuclei. If the galaxy does not contain a massive primordial black hole, the first cycle of disc evolution will lead to the initial formation of a rapidly rotating self-gravitating disc-like configuration which, it was assumed, would resemble a low-entropy spinar in appearance and properties. Non-thermal 'magnetic dipole' radiation from this body might in principle explain the most luminous quasars (even if the spinar only had a mass $10^4 - 10^6$ $M_\odot$), but its late evolution, crucially dependent on the effect of the onset of nuclear burning, could not be predicted reliably because detailed models are not yet available. Three possible evolutionary end-points were distinguished. In the first, the body collapsed into a black hole; subsequent cycles of activity had then to be due to processes of accretion. Although in this case the structure and evolution of the disc during second and subsequent cycles would still be similar to that described here, significant quantitative differences might occur close to the centre. In the second possible kind of evolution, the body evolved 'quasi-statically' until the onset of explosive nuclear burning. In this case nuclear reactions led to a disruptive nuclear explosion, and if this occurred when the spinar's binding energy was sufficiently high (but still less than $E_{nuc} \sim 0{\cdot}7\ \%\ Mc^2$) it <u>might</u> be possible to explain observed nuclear activity as due to the repeated formation and explosive disruption of a low-mass spinar. The third possibility - in which the body does undergo gravitational collapse, but not into a black hole - was not discussed here in any detail, because it is then no longer clear just what input physics to use to describe the body's subsequent evolution. Clube (1978) has argued strongly in favour of this possibility, and has suggested the temporary action of a greatly increased gravitational attraction ('supermass') to explain a number of detailed observations of our own Galaxy, in particular the apparent outward motion of the local standard of rest ($\Pi \sim 40$ km s^{-1}). In view of this possibility, it is extremely important firstly to decide whether or not the evolution of a low-mass spinar does lead to gravitational collapse, and secondly (if it does), to determine the detailed evolutionary history of a hypothetical black hole in the nucleus in order to make a reliable estimate of its present-day mass.

Observations of our Galaxy were reviewed in Chapter VII. These were shown to be quite understandable in terms of the present theory provided that the Galaxy has recently finished a period of nuclear activity. The presence of large quantities of gas in our Galaxy's nuclear region may be interpreted as supporting evidence for the ideas presented here, and the apparent absence of a supermassive black hole ($M \gtrsim 5 \times 10^6\ M_\odot$) argues against the black hole theory. It was tentatively concluded that black holes do not explain observed nuclear activity, although which of the two other possible evolutionary tracks actually occurs in nature can not be determined until further more detailed investigations have been made.

In summary, the present thesis has shown (1) that recurrent nuclear activity can be understood as due to the evolution of stellar mass loss in normal galactic nuclei;. (2) that an important intermediary leading to the formation of an active nucleus is a dense, large-scale cold nuclear disc, composed of gas predominantly in molecular form; and (3) that nuclear activity is best explained by the formation and subsequent evolution of a low-mass spinar ($M \lesssim 10^6\ M_\odot$) which evolves ultimately towards total destruction either by a disruptive nuclear explosion, or, from a state more tightly bound than E_{nuc}, by processes involving as yet undiscovered physical laws. It is not yet possible to regard these conclusions as completely firm, since they are too heavily dependent on arguments which themselves involve uncertain areas of astronomy, such as star formation and the detailed evolution of molecular clouds and H II regions, but it is important to emphasise the overall coherence of the argument. Much more work is necessary before the proposed evolution can be finally accepted: Firstly, a detailed study of the evolution and structure of large-scale massive nuclear discs is required, not only to confirm the order of magnitude estimates obtained in Chapter V but also to determine under what circumstances such discs might become observable in pre-active galaxies. Secondly, it is important to make a detailed study of the evolution of the central parts of the disc following the onset of the first gravitational instability. The rate of inward transport due to viscous processes is crucially dependent on the disc's structure during this phase of evolution, and this in turn determines the pre-conditions for the formation of the active nucleus and its environment during its later evolution. It is also important to determine the initial mass function of stars formed in the central parts of the disc, since this will affect the prediction of observed M/L ratios in the (seeing-limited) nuclear regions of distant galaxies: Confirmation of nuclear evolution of the kind suggested here might be possible by comparing observed nuclear M/L ratios with those predicted by theory (cf. observations of M 87; Sargent et al, 1978; Young et al, 1978). Thirdly, in view of the theoretical arguments presented in this thesis, a more detailed study of the observational evidence for recurrence (cf. Chapter II) would be useful; especially with regard to the question of how the mean interval between active periods depends on observational parameters such as galaxy type, mass, and net rotational velocity. Lastly, if the arguments presented here <u>are</u> accepted, the most important theoretical problem to have been isolated concerns the detailed formation and evolution of low-mass spinars in normal galactic nuclei. If such objects do undergo gravitational collapse, the body either forms a massive black hole or enters a phase of evol-

APPENDIX : ESTIMATE OF THE KINEMATIC VISCOSITY

Any estimate of the kinematic viscosity in the nuclear disc must be based not only on a detailed physical model of the disc's structure, but also on a picture as to how randomly-moving elements of the disc actually transport momentum from one part of the system to another. In order to try and bracket the likely range of possibilities, estimates of the kinematic viscosity are given here for two idealised disc models. In the first, the disc is pictured as composed entirely of individual particles (i.e. molecules or 'clouds'); whereas in the second, it is assumed that the disc is turbulent, with a viscosity dominated by the motions of individual turbulent elements or 'eddies'. Although a realistic disc model will probably be more complicated than either of these two extremes, it seems reasonable to expect that a more detailed estimate of the viscosity would lead to a value for ν of the same order of magnitude as the estimates obtained here. Evolution due to possible magnetic torques is not considered, but it may be argued qualitatively (cf. pages 165 - 166) that the magnetic 'viscosity' is of the same order as the turbulent viscosity. Inclusion of magnetic effects may therefore lead to an additional term in the final expression for ν of the same order as ν_{turb}

First consider the case in which the disc is composed entirely of individual particles (either molecules or clouds) with space density n, mass m, mean free path λ and mean velocity $\bar{c}$. The viscosity of this model will be referred to as ν_{part} , and may be determined by direct application of standard gas kinetic theory. In a finite system, particles arriving at a given point can come from distances no larger than R_{max} (say), and the kinematic viscosity may be shown to be

$$\nu_{part} \sim f \bar{c} \lambda \left\{ 1 - e^{-\frac{R_{max}}{\lambda}} \left(1 + \frac{R_{max}}{\lambda} \right) \right\} \qquad (A.1)$$

where f is a factor of order unity whose precise value depends on the assumptions made as to the velocity distribution. For an isotropic velocity distribution $f = \frac{1}{3}$, whereas for a two-dimensional system (i.e. velocities restricted to a plane) $f = \frac{1}{2}$.

In the present problem R_{max} may be estimated to be roughly equal to the amplitude of the epicyclic motions in the plane. This gives

$$R_{max} \sim \bar{c}/\kappa$$

(A.1) may therefore be re-written in the form

$$\nu_{part} \sim \frac{1}{3} \frac{\bar{c}^2}{\kappa} \times g \qquad (A.2)$$

where the function g is defined by

$$g = 3f \left(\frac{\lambda}{R_{max}} \right) \left\{ 1 - e^{-\frac{R_{max}}{\lambda}} \left(1 + \frac{R_{max}}{\lambda} \right) \right\} \qquad (A.3)$$

In the case in which the particles are molecules, it may easily be verified that for typical disc densities (Section V-4) $\lambda \ll R_{max}$. In this case

$$\nu(\lambda \ll R_{max}) \sim f \bar{c} \lambda$$

which, for $f = \frac{1}{3}$, is the usual formula for the molecular kinematic viscosity. In this limit (A.3) becomes

$$g(\lambda \ll R_{max}) \sim 3f \times \left(\frac{\lambda}{R_{max}} \right) \ll 1 \qquad (A.4)$$

On the other hand, if the disc consists of a number of small dense clouds, it is possible that $\lambda \gg R_{max}$. In this limit (A.1) becomes

$$\nu(\lambda \gg R_{max}) \sim f\,\bar{c}\,\frac{R_{max}^2}{\lambda}$$

which implies

$$g(\lambda \gg R_{max}) \sim 3f_*\left(\frac{R_{max}}{\lambda}\right) \ll 1 \qquad (A.5)$$

Thus, the kinematic viscosity due to particle collisions will be small not only when the mean free path is very short, but also when it is very long.(For example compare molecular viscosity with viscosity in a stellar disc). Further progress requires information as to the value of the mean free path. On the assumption that the disc is composed entirely of spherical clouds of radius a and mean density ρ_{cl}, it may be shown that the mean free path $\lambda \sim 1/(4\pi a^2 n)$ reduces to

$$\lambda \sim a/3f' \qquad (A.6)$$

where $f' \equiv \rho/\rho_{cl}$ is the filling factor and $\rho = n\,m$ is the mean disc density. If the clouds are assumed to be composed of cool gas ($T \sim 100 - 1000$ K, say) in rough pressure balance with a hot intercloud medium ($T \sim 10^4$ K, say), the filling factor should lie in the range $0{\cdot}01 \lesssim f' \lesssim 0{\cdot}1$; a result that is in good agreement with observations of our own Galaxy (e.g. Lynden-Bell & Pringle, 1974; Mezger, 1974). On this model the r.m.s. gas velocity within a cloud (V_{cl}, say) is less than or equal to the mean random velocity of the clouds themselves ($\bar{c}$), which implies that the cloud radii a should be less than the disc's semi-width multiplied by $V_{cl}/\bar{c}$. For a non self-gravitating disc, the width is given by (36), so

$$\alpha \lesssim \frac{1}{2} H(\varpi) \times \frac{V_{cl}}{\bar{c}} \sim \left(\frac{\bar{c}}{\kappa}\right) \times \left(\frac{V_{cl}}{\bar{c}}\right) \qquad (A.7)$$

where κ has been assumed equal to $\sqrt{2}\, V_o / \varpi$. Thus

$$\lambda \lesssim \frac{1}{3}\left(\frac{V_{cl}}{\bar{c}}\right) \cdot \frac{R_{max}}{f'}$$

which, using the approximate relation $f' \sim \left(\frac{V_{cl}}{\bar{c}}\right)^2$, gives

$$\lambda \lesssim \frac{1}{3} f'^{-1/2} R_{max} \qquad (A.8)$$

Substituting this into (A.3) therefore gives for the stated range of f':

$$0{\cdot}4 f \lesssim g \lesssim 0{\cdot}8 f \qquad (A.9)$$

The kinematic viscosity of a cloudy disc may therefore be written in the form

$$\nu_{cl} \sim \frac{1}{3} \frac{\bar{c}^2}{\kappa} \times g_{cl}$$

where g_{cl} is typically of the order of $0{\cdot}2$. Compared with this, molecular viscosity (e.g. equation (A.4)) is completely negligible.

We now estimate the kinematic viscosity in a model in which the disc is assumed to be composed of gas in a state of fully developed stationary turbulence. If $\rho.F(k)dk$ is defined to be the energy per unit volume contained in turbulent eddies with wave numbers between k and $k+dk$, the turbulent kinematic viscosity due to eddies with wave numbers larger than k may be written (e.g. Heisenberg, 1948; Chandrasekhar, 1949) in the form

$$\nu_k = \kappa \int_k^\infty \left(\frac{F(k)}{k^3}\right)^{1/2} dk \qquad (A.11)$$

where K is a numerical coefficient of order unity. Experimental data seem to indicate $K \sim 0{\cdot}8$ (Heisenberg, 1948). If the spectrum of the turbulence is stationary, and the molecular viscosity is very small (so that the system has a very large Reynolds number), it may be shown (e.g. Chandrasekhar, 1949) that F(k) reduces to the equilibrium Kolmogoroff spectrum; i.e.

$$F(k) = F(k_0)\left(\frac{k_0}{k}\right)^{5/3} \qquad k \geqslant k_0 \qquad (A.12)$$

where k_0 is the wave number of the largest turbulent eddies. The turbulent viscosity is given by ν_{k_0}, so substituting (A.12) into (A.11) thus implies

$$\nu_{turb} \sim \frac{3K}{4}\left(\frac{F(k_0)}{k_0}\right)^{1/2} \qquad (A.13)$$

The turbulent energy density is given by

$$\frac{1}{2}\rho\,\bar{c}^2 \sim \rho\int_{k_0}^{\infty} F(k)\,dk$$

$$\Rightarrow \bar{c}^2 \sim 3.F(k_0).k_0$$

$$\Rightarrow \nu_{turb} \sim \frac{\sqrt{3}\,K}{4}\,\frac{\bar{c}}{k_0} \qquad (A.14)$$

Heisenberg defines the diameter of the largest turbulent eddy to be $\lambda_0 = \pi/k_0$, so

$$\nu_{turb} \sim \frac{\sqrt{3}\,K}{4\pi}\,\bar{c}\,\lambda_0$$

$$\text{i.e}\quad \nu_{turb} \sim 0{\cdot}1\,\bar{c}\,\lambda_0 \qquad (A.15)$$

where it has been assumed that $K \sim 0{\cdot}8$. Supersonic motions are quickly dissipated in a dense medium, so the appropriate value of $\bar{c}$ will be on the order of the sound speed.

It remains now to estimate the diameter λ_o of the largest turbulent eddies. As noted by Heisenberg (1948) and Cameron (1970) this is essentially a geometrical (as opposed to statistical) problem, since the precise value of λ_o must depend primarily on specific details of the particular system under investigation and on the detailed physical processes which are the underlying source of turbulent kinetic energy. In the present problem, prior to the onset of gravitational instability turbulence will be driven by both infalling material and the passage of stars through the disc. The relevant length-scale on which turbulence will be generated is therefore on the order of the disc width, which (by equation (36)) is of order $2\ R_{max}$. During this phase of evolution the turbulent viscosity is thus given approximately by

$$\nu_{turb} \sim \frac{1}{3}\ \frac{\bar{c}^2}{\Pi} \times g_{turb}$$

where $g_{turb} \sim 3/5$; i.e. of order unity.

When the disc becomes gravitationally unstable, turbulent eddies with diameters on the order of the critical unstable wavelength are likely to form, which implies (cf. equation (108))

$$\lambda_o \sim \lambda_{crit} \sim \frac{2V_s^2}{G\sigma_{crit}} \qquad (A.16)$$

Substituting for σ_{crit} (equation (29)) and assuming $V_s \sim \sqrt{2}\ \bar{c}$ thus implies $\lambda_o \sim 2\sqrt{2}\,\pi\,\ell_{max}$, which implies $g_{turb} \sim 3$. As was noted at the beginning of the Appendix, inclusion of possible effects due to magnetic fields might increase the effective g to a value somewhat above this (i.e. to $g \sim 5$, say).

In conclusion, it has been shown that the kinematic viscosity of the disc during both its main phases of evolution may be written in the form

$$\nu \sim \frac{1}{3} \frac{\bar{c}^2}{\kappa} \times g \tag{A.17}$$

where $\bar{c}$ is on the order of the sound speed, κ is the epicyclic frequency, and g a numerical factor which depends on the detailed physical model of the disc. During the phase of evolution which precedes the onset of gravitational instability, the disc will be composed of gas distributed homogeneously, and viscosity will be caused by molecular viscosity or turbulent motions. In the absence of turbulence it was shown that $g_{mol} \ll 1$, indicating that molecular viscosity is negligibly small. However such a system would have a very large Reynolds number, and for this reason and also because infall and the passage of stars through the disc are both expected to drive some mass-motions, the gravitationally stable disc is expected to be in a state of turbulence. On the assumption that turbulence is driven with a scale-size of the same order as the disc width, g_{turb} was shown to be of order unity. After the onset of gravitational instability the physical state of the disc becomes extremely complex, and it is not easy to decide which of the two extreme disc models considered here is the more realistic. In the case of a cloudy disc g_{cl} was shown to be of order $0{\cdot}2$; but in the alternative case of a turbulent system containing turbulent eddies with sizes up to that of the critical unstable wavelength λ_{crit}, it was shown that $g_{turb} \sim 3$. Inclusion of magnetic 'viscosity' might lead to an additional torque of the same order as that due to turbulence. None of the values of g calculated here should be regarded as more accurate than a factor of about two, and for this reason it is assumed in the present thesis that throughout the disc's evolution the kinematic viscosity may be approximated by taking $g \sim 1$.

ACKNOWLEDGMENTS

I should like to express my sincere thanks to my two supervisors, Dr P.W.J.L. Brand and Dr S.V.M. Clube, with whom I have spent many hours of fruitful discussion, and who have greatly assisted me during the course of this project by always allowing me to make free use of their time. In particular, discussions with Peter have been especially useful in providing many insights into the detailed structure and evolution of the disc; and discussions with Victor have greatly helped in keeping the fundamental aims of the project to the fore, always forcing me to regard the thesis as a whole. Between them they have taught me a great deal of astronomy - in particular how to do astronomical research - and their judgement has to a large extent helped to shape the present thesis and keep the project on course during its three years' evolution.

During the progress of this project I have also benefitted by discussions with a large number of astronomers other than my supervisors, and I should like to thank Dr D.C. Heggie and Professor L. Mestel for helpful and encouraging comments during early phases of the project. Colleagues at Edinburgh, both staff and students, have also helped by accepting uncomplainingly my frequent interruptions and requests for information.

Finally, for my financial support during the last three years, I thank the Cormack Committee of the Royal Society of Edinburgh for providing funds through the Robert Cormack Bequest.

REFERENCES

Adams, T.F. (1977) 'A survey of Seyfert galaxies based on large-scale image tube plates' Astrophys J (Suppl) 33 19 1977

Allen, C.W. (1973) 'Astrophysical Quantities' Athlone Press, Univ of London 1973

Ambartsumian, V.A. (1958) 'On the evolution of galaxies' 11 th Solvay Conf.: Structure and Evolution of the Universe; p.241, 1958

Ambartsumian, V.A. (1961) 'Instability phenomena in systems of galaxies' Astron J 66 536 1961

Ambartsumian, V.A. (1965) 'On the nuclei of galaxies and their activity' 13th Solvay Conf.: Structure and evolution of galaxies; p.1, 1965

Ambartsumian, V.A. (1971) 'Introduction' Nuclei of Galaxies; p.1, 1971

Andriesse, C.D. (1974) 'Far infrared properties of interstellar grains' Astron & Astrophys 37 257 1974

Andriesse, C.D. & de Vries, J. (1976) 'Distribution of stars and dust near the Galactic Centre derived from far infrared observations' Astron & Astrophys 50 99 1976

Appenzeller, I. & Tscharnuter, W. (1973) 'Exploding massive objects' Astron & Astrophys 25 125 1973

Arons, J., Kulsrud, R.D. & Ostriker, J.P. (1975) 'A multiple pulsar model for quasi-stellar objects and active galactic nuclei' Astrophys J 198 687 1975

Baade, W. & Minkowski, R. (1954) 'Identification of the radio sources in Cassiopeia, Cygnus A, and Puppis A' Astrophys J 119 206 1954

Balick, B., & Brown, R.L. (1974) 'Intense sub-arcsecond structure in the Galactic Centre' Astrophys J 194 265 1974

Balick, B., & Sanders, R.H. (1974) 'Radio fine structure in the Galactic Centre' Astrophys J 191 325 1974

Becklin, E.E, Matthews, K., Neugebauer, G., & Willner S.P. (1978) 'Infrared observations of the Galactic Centre I. Nature of the compact sources' Astrophys J 219 121 1978

Becklin, E.E., & Neugebauer, G. (1968) 'Infrared observations of the Galactic Centre' Astrophys J 151 145 1968

Becklin, E.E., & Neugebauer, G. (1975) 'High resolution maps of the Galactic Centre at 2·2 and 10 microns' Astrophys J (Lett) 200 L71 1975

Bisnovatyi-Kogan, G.S., & Blinnikov, S.I. (1977) 'Disk accretion onto a black hole at subcritical luminosity' Astron & Astrophys 59 111 1977

Bohuski, T.J. (1973) 'Structure of the H II regions M 8 and M 20 II. Electron densities' Astrophys J 184 93 1973

Burbidge, G.R. (1956) 'On synchrotron radiation from Messier 87' Astrophys J 124 416 1956

Burbidge, G.R. (1970) 'The nuclei of galaxies' Ann Rev Astron & Astrophys 8 369 1970

Burbidge, G.R. (1971) 'Theoretical considerations regarding non-thermal emission and ejection of matter from galactic nuclei' Nuclei of Galaxies p.411, 1971

Burbidge, G.R. & Hoyle, F. (1963) 'The Galactic Halo – a transient phenomenon' Astrophys J 138 57 1963

Burke, J.R. & Silk, J. (1974) 'Dust grains in a hot gas I. Basic physics' Astrophys J 190 1 1974

Callahan, P.S. (1977) 'On the accretion disk model of QSOs' Astron & Astrophys 59 127 1977

Cameron, A.G.W. (1962) 'Star formation in elliptical galaxies and intense radio sources' Nature 194 963 1962

Cameron, A.G.W. (1969a) 'The pre-Hayashi phase of stellar evolution' Low Luminosity Stars (Ed: S.S. Kumar) p.423, 1969

Cameron, A.G.W. (1969b) 'Physical conditions in the primitive solar nebula' Meteorite Research (Ed: P.M. Millman) p.7 1969

Cameron, A.G.W. (1970) 'Formation of the Earth-Moon system' Trans Amer Geophys Union 51 628 1970

Capriotti, E.R. (1973) 'The structure and evolution of planetary nebulae' Astrophys J 179 495 1973

Castor, J., McCray, R., & Weaver, R. (1975) 'Interstellar bubbles' Astrophys J (Lett) 200 L107 1975

Chandrasekhar, S. (1949) 'On Heisenberg's elementary theory of turbulence' Proc Roy Soc (Lond) 200A 20 1949

Clube, S.V.M. (1978) 'Does our Galaxy have a violent history ?' Vistas in Astronomy 22 77 1978

Cohen, R.J. (1975) 'Neutral hydrogen in the Galactic Centre region I. The observations' Mon Not Roy Astron Soc 171 659 1975

Cohen, R.J., & Davies, R.D. (1976) 'Neutral hydrogen in the Galactic Centre region II. Location of the emission features' Mon Not Roy Astron Soc 175 1 1976

Cohen, R.J., & Davies, R.D. (1978) 'Neutral hydrogen in the Galactic Centre region IV. The Galactic Nucleus' Preprint

Cohen, R.J., & Few, R.W. (1976) 'Observations of OH in the Galactic Centre' Mon Not Roy Astron Soc 176 495 1976

Colla, G., Fanti, R., Gioia, I., Lari, C., Lequeux, J., Lucas, R., & Ulrich, M-H (1975) 'A complete survey of radio sources identified with elliptical galaxies: Radio luminosity function and other properties' Astron & Astrophys 38 209 1975

Cox, D.P., & Smith, B.W. (1974) 'Large-scale effects of supernova remnants on the Galaxy: Generation and maintenance of a hot network of tunnels' Astrophys J (Lett) 189 L105 1974

Cox, J.P., & Guili, R.T. (1968) 'Principles of stellar structure' Gordon & Breach, 1968. Ch. 16 Stellar Opacity p.353

Dalgarno, A., & McCray, R.A. (1972) 'Heating and ionisation of H I regions' Ann Rev Astron & Astrophys 10 375 1972

Davies, R.D., Walsh, D., & Booth, R.S. (1976) 'The radio source at the Galactic Nucleus' Mon Not Roy Astron Soc 177 319 1976

De Bruyn, A.G. (1977) 'A radio continuum study of four spiral galaxies with an unusual radio morphology' Astron & Astrophys 58 221 1977

De Bruyn, A.G., & Wilson, A.S. (1978) 'The radio properties of Seyfert galaxies' Astron & Astrophys 64 433 1978

De Vaucouleurs, G. & de Vaucouleurs, A. (1968) 'Photographic, photometric, and spectroscopic observations of Seyfert galaxies' Astron J 73 858 1968

De Young, D.S. (1976) 'Extended extragalactic radio sources' Ann Rev Astron & Astrophys 14 447 1976

Dokuchaev, V.I., & Ozernoy, L.M. (1977) 'Tidal disruption of stars and the evolution of a massive black hole under the conditions of the Galactic Centre' Sov Astron Lett 3 209 1977

Downes, D. & Martin, A.H.M. (1971) 'Compact radio sources in the Galactic Nucleus' Nature 233 112 1971

Eardley, D.M. (1974) 'Death of white holes in the early universe' Phys Rev Lett 33 442 1974

Ekers, R.D. (1975) 'Radio continuum emission from galaxies' Structure and Evolution of Galaxies (Ed: G. Setti) p. 217, 1975

Ekers, R.D. & Lynden-Bell, D. (1971) 'High resolution observations of the Galactic Centre at 5 GHz' Astrophys Lett 9 189 1971

Elmegreen, B.G. & Lada, C.J. (1977) 'Sequential formation of subgroups in OB associations' Astrophys J 214 725 1977

Faber, S.M. & Gallagher, J.S. (1976) 'H I in early-type galaxies II. Mass loss and galactic winds' Astrophys J 204 365 1976

Falle, S.A.E.G. (1975) 'Numerical calculations of thin shell formation in supernova remnants' Mon Not Roy Astron Soc 172 55 1975

Field, G.B. (1964) 'Quasi-stellar radio sources as spherical galaxies in the process of formation' Astrophys J 140 1434 1964

Flasar, F.M. & Morrison, P (1976) 'A spinar model of Cygnus A' Astrophys J 204 352 1976

Fowler, W.A. (1966) 'Supermassive stars, quasars, and extragalactic radio sources' High Energy Astrophysics (Enrico Fermi School of Physics) p.313, 1966

Fowler, W.A. (1971) 'Periodicity and luminosity of the 'pulsar' model of quasars' I.A.U. Symp 46 364 1971

Frank, J. (1976) 'Supermassive black holes dethroned ?' Observatory 96 198 1976

Frank, J. & Rees, M.J. (1976) 'Effects of massive central black holes on dense stellar systems' Mon Not Roy Astron Soc 176 633 1976

Fricke, K.J. (1973) 'Dynamical phases of supermassive stars' Astrophys J 183 941 1973

Fricke, K.J. (1974) 'Dynamical phases of rotating supermassive stars' Astrophys J 189 535 1974

Gatley, I., Becklin, E.E., Werner, M.W., & Wynn-Williams, G.C. (1977) 'Airborne far-infrared observations of the Galactic Centre Region' Astrophys J 216 277 1977

Ginzburg, V.L. (1960) 'The nature of radio galaxies' Sov Astron 5 282 1961

Gisler, G.R. (1976) 'The fate of gas in elliptical galaxies and the density evolution of radio sources' Astron & Astrophys 51 137 1976

Goldreich, P. & Lynden-Bell, D. (1965a) 'Gravitational instabilities of uniformly rotating discs' Mon Not Roy Astron Soc 130 97 1965

Goldreich, P. & Lynden-Bell, D. (1965b) 'Spiral arms as sheared gravitational instabilities' Mon Not Roy Astron Soc 130 125 1965

Goldreich, P. & Ward, W.R. (1973) 'The formation of planetisimals' Astrophys J 183 1051 1973

Goldsmith, P.F. & Langer, W.D. (1978) 'Molecular cooling and thermal balance in dense interstellar clouds' Astrophys J 222 881 1978

Gott III, J.R. & Thuan, T.X. (1976) 'On the formation of spiral and elliptical galaxies' Astrophys J 204 649 1976

Greenberg, J.M. (1971) 'interstellar grain temperatures. Effects of grain materials and radiation fields' Astron & Astrophys 12 240 1971

Hara, T. & Ikeuchi, S (1976) 'Formation and evolution of a high-density gaseous disc in a dense stellar system' Prog Theor Phys (Japan) 56 531 1976

Harvey, P.M., Campbell, M.F., & Hoffmann, W.F. (1976) 'High-resolution far infra-red observations of the Galactic Centre' Astrophys J (Lett) 205 L69 1976

Heisenberg, W. (1948) 'On the theory of statistical and isotropic turbulence' Proc Roy Soc (Lond) 195A 402 1948

Hills, J.G. (1975) 'Possible power source of Seyfert galaxies and QSOs' Nature 254 295 1975

Hjellming, R.M. (1971) 'Black and white holes' Nature (Phys Sci) 231 20 1971

Hoffmann, W.F., Frederick, C.L., & Emery, R.J. (1971) '100-micron map of the Galactic Centre region' Astrophys J (Lett) 164 L23 1971

Hoyle, F. (1961) 'Galactic flares' Observatory 81 39 1961

Hoyle, F. (1971) 'On the nature of compact objects' Nuclei of Galaxies p. 583, 1971

Hoyle, F. & Fowler, W.A. (1963) 'On the nature of strong radio sources' Mon Not Roy Astron Soc 125 169 1963

Hoyle, F. & Fowler, W.A. (1973) 'On the origin of deuterium' Nature 241 384 1973

Hoyle, F. & Narlikar, J.V. (1966a) 'On the effects of non-conservation of baryons in cosmology' Proc Roy Soc (Lond) 290A 143 1966

Hoyle, F. & Narlikar, J.V. (1966b) 'A radical departure from the 'steady-state' concept in cosmology' Proc Roy Soc (Lond) 290A 162 1966

Hoyle, F. & Narlikar, J.V. (1966c) 'On the formation of elliptical galaxies' Proc Roy Soc (Lond) 290A 177 1966

Huchra, J.P. (1977) 'The nature of Markarian galaxies' Astrophys J (Suppl) 35 171 1977

Huchra, J.P. & Sargent, W.L.W. (1973) 'The space density of the Markarian galaxies including a region of the South Galactic Hemisphere' Astrophys J 186 443 1973

Ito, K., Matsumoto, T., & Uyama, K. (1977) 'Infrared profile of central region of our Galaxy at 2·47 μm' Nature 265 517 1977

Jeans, J. (1928) 'Astronomy & Cosmology' p. 351, 1928

Johnson, H.E. & Axford, W.I. (1971) 'Galactic Winds' Astrophys J 165 381 1971

Kaifu, N., Iguchi, T., & Kato, T. (1974) 'Gaseous structure in the Galactic Centre region' Publ Astron Soc Japan 26 117 1974

Kaifu, N., Kato, T., & Iguchi, T. (1972) '270 pc expanding ring at the Galactic Centre' Nature (Phys Sci) 238 105 1972

Kaifu, N., Morris, M., Palmer, P., & Zuckerman, B (1975) 'Distribution of NH_3 clouds in the Galactic Centre region' Astrophys J 201 98 1975

Kellermann, K.I., Shaffer, D.B., Clark, B.G., & Geldzahler, B.J. (1977) 'The small radio source at the Galactic Centre' Astrophys J (Lett) 214 L61 1977

Kellermann, K.I., Shaffer, D.B., Pauliny-Toth, I.I.K., Preuss, E., & Witzel, A. (1976) 'Observations of a radio source in the nucleus of M 81 with dimensions less than 1300 AU' Astrophys J (Lett) 210 L121 1976

Khachikian, E. Ya., & Weedman, D.W. (1974) 'An atlas of Seyfert galaxies' Astrophys J 192 581 1974

Krugel, E., & Tutukov, A.V. (1978) 'A model for the radiation from the Galactic Centre' Astron & Astrophys 63 375 1978

Lang, K.R., & Terzian, Y. (1969) 'Radio emission of normal galaxies' Astrophys Lett 3 29 1969

Larson, R.B. (1973) 'A simple probabilistic theory of fragmentation' Mon Not Roy Astron Soc 161 133 1973

Larson, R.B. (1974a) 'Dynamical models for the formation and evolution of galaxies' Mon Not Roy Astron Soc 166 585 1974

Larson, R.B. (1974b) 'Effects of supernovae on the early evolution of galaxies' Mon Not Roy Astron Soc 169 229 1974

Layzer, D (1965) 'Magnetoturbulence and astronomy I. The nature and origin of non-thermal radio sources' Astrophys J 141 837 1965

Leung, C.M. (1975) 'Radiation transport in dense interstellar dust clouds I. Grain temperature' Astrophys J 199 340 1975

Lewis, B.M. (1972) 'Seyfert galaxies' I.A.U. Symp 44 267 1972

Light, E.S., Danielson, R.E., & Schwartzschild, M. (1974) 'The nucleus of M31' Astrophys J 194 257 1974

Limber, D.N. (1960) 'The universality of the initial luminosity function' Astrophys J 131 168 1960

Lo, K.Y., Schilizzi, R.T., Cohen, M.H., & Ross, H.N. (1975) 'VLBI observations of the compact radio source in the centre of our Galaxy' Astrophys J (Lett) 202 L63 1975

Lynden-Bell, D (1969) 'Galactic nuclei as collapsed old quasars' Nature 223 690 1969

Lynden-Bell, D (1971) 'Formation and evolution of bright black holes' Nuclei of Galaxies p.527 1971

Lynden-Bell, D. & Pringle, J.E. (1974) 'The evolution of viscous discs and the origin of the nebular variables' Mon Not Roy Astron Soc 168 603 1974

Lynden-Bell, D. & Rees, M.J. (1971) 'On quasars, dust and the Galactic Centre' Mon Not Roy Astron Soc 152 461 1971

Markarian, B.E. (1969) 'Galaxies with ultraviolet continuum III' Astrophysics 5 286 1969

Mathews, W.G. (1972) 'Collapse in galactic nuclei' Astrophys J 174 101 1972

Mathews, W.G. & Baker, J.C. (1971) 'Galactic winds' Astrophys J 170 241 1971

Matsuda, T. & Nelson, A.H. (1977) 'A Galactic vacuum cleaner ?' Nature 266 607 1977

McCrea, W.H. (1964) 'Continual creation' Mon Not Roy Astron Soc 128 335 1964

Mestel, L. (1963) 'On the Galactic law of rotation' Mon Not Roy Astron Soc 126 553 1963

Mestel, L. (1965) 'Problems of star formation - I, II' Quat J Roy Astron Soc 6 161, 265 1965

Mezger, P.G. (1974) 'The Centre of the Galaxy' Proc. ESO/SRC/CERN Conf. on Research Prog. for new large telescopes p.79 1974

Mezger, P.G., Churchwell, E.B., & Pauls, T. (1974) 'Ionised gas in the direction of the Galactic Centre I. Kinematics and physical conditions in the Nuclear Disc' Stars and the Milky Way System (Ed: L.N. Mavridis) p. 140, 1974

Mills, B.Y. (1956) 'The radio source near the Galactic Centre' Observatory 76 65 1956

Moorwood, A.F.M. (1974) 'H II regions and the Galactic Centre' Proc. 8th ESLAB Symp. ESRO SP-105

Morgan, W.W., Walborn, N.R. & Tapscott, J.W. (1971) 'An optical form morphology of Seyfert galaxies' Nuclei of Galaxies p.27 1971

Morton, D.C., & Elmegreen, B.G. (1976) 'Velocity dispersions in galaxies V. The nuclei of M 31 and M 32' Astrophys J 205 63 1976

Ne'eman, Y. (1965) 'Expansion as an energy source in quasi-stellar radio sources' Astrophys J 141 1303 1965

Neugebauer, G., Becklin, E.E., Matthews, K., & Wynn-Williams, C.G. (1978) 'Infrared observations of the Galactic Centre III. 2·2 μm spectroscopy' Astrophys J 220 149 1978

Novikov, I.D. (1964) 'Delayed explosion of a part of the Friedman universe, and quasars' Sov Astron 8 857 1965

Oort, J.H. (1971) 'Composition and activity of the Nucleus of our Galaxy and comparison with M 31' Nuclei of Galaxies p.321 1971

Oort, J.H. (1974a) 'The Galactic Centre' I.A.U. Symp. 60 539 1974

Oort, J.H. (1974b) 'Recent radio work in nearby galaxies' Galaxies and Relativistic Astrophysics (Ed: Barbanis & Hadjidemetriou) p.1 1974

Oort, J.H. (1977) 'The Galactic Centre' Ann Rev Astron & Astrophys 15 295 1977

Oort, J.H. (1978) 'Eruptive phenomena near the Galactic Centre' Physica Scripta 17 175 1978

Opher, R. (1974) 'A theory of galactic nuclei and quasi-stellar objects' Astrophys J 188 201 1974

Ozernoy, L.M. (1968) 'The magnetohydrodynamic model of quasars' Highlights in Astronomy 1 384 1968

Ozernoy, L.M. (1974) 'Galactic nuclei' Proc. of First European Astron. Meeting 1972 ; 3 65 1974 (Ed: Barbanis & Hadjidemetriou)

Ozernoy, L.M., & Usov, V.V. (1971) 'Structure and evolution of supermassive rotating magnetic polytropes' Astrophys & Space Science 13 3 1971

Ozernoy, L.M., & Usov, V.V. (1973) 'Supermassive oblique rotator: Electrodynamics, evolution, observational tests' Astrophys & Space Science 25 149 1973

Ozernoy, L.M., & Usov, V.V. (1977) 'Regular optical variability of quasars and nuclei of galaxies as a clue to the nature of their activity' Astron & Astrophys 56 163 1977

Pastoriza, M.G. (1975) 'Studies of galaxies with peculiar nuclei' Astrophys & Space Science 33 173 1975

Pauls, T., Mezger, P.G., & Churchwell, E.B. (1974) 'Recombination line emission from an H II region surrounding the nucleus of the Galaxy' Astron & Astrophys 34 327 1974

Pauls, T., Downes, D., Mezger, P.G., & Churchwell, E (1976) 'Recombination line and continuum observations of the Galactic Centre at 10 GHz' Astron & Astrophys 46 407 1976

Peters III, W.L. (1975) 'Models for the inner regions of the Galaxy I. An elliptical streamline model' Astrophys J 195 617 1975

Piddington, J.H. (1964) 'The magnetic fields and radio emission of galaxies' Mon Not Roy Astron Soc 128 345 1964

Piddington, J.H. (1966) 'Galactic explosions, radio galaxies and quasi-stellar sources' Mon Not Roy Astron Soc 133 163 1966

Pringle, J.E., Rees, M.J., & Pacholczyk, A.G. (1973) 'Accretion onto massive black holes' Astron & Astrophys 29 179 1973

Readhead, A.C.S., Cohen, M.H., & Blandford, R.D. (1978) 'A jet in the nucleus of NGC 6251' Nature 272 131 1978

Rees, M.J. (1974) 'Astrophysical implications of extragalactic radio sources' Proc. First Europ. Astron. Meeting, 1972 3 190 1974. (Ed: Barbanis & Hadjidemetriou)

Rees, M.J. (1977) 'Quasar theories' Ann New York Acad Sci 302 613 1977 (Proc. 8th Texas Symp on Rel. Astrophys)

Rees, M.J. (1978) 'Accretion and the quasar phenomenon' Physica Scripta 17 193 1978

Richstone, D.O., & Morton, D.C. (1975) 'Velocity dispersions in Galaxies IV: The nucleus of NGC 1068' Astrophys J 201 289 1975

Rickard, L.J., Palmer, P., Morris, M., Turner, B.E., & Zuckerman, B (1977) 'Extragalactic carbon monoxide' Astrophys J 213 673 1977

Rieke, G.H., Grasdalen, G.L., Kinman, T.D., Hintzen, P., Wills, B.J., & Wills, D. (1976) 'Photometric and spectroscopic observations of the BL Lac object AO 0235+164' Nature 260 754 1976

Rieke, G.H., Telesco, C.M., & Harper, D.A. (1978) 'The infrared emission of the Galactic Centre' Astrophys J 220 556 1978

Rougoor, G.W. (1964) 'The neutral hydrogen in the central region of the Galactic System' Bull Astron Inst Neth 17 381 1964

Rougoor, G.L. & Oort, J.H. (1960) 'Distribution and motion of interstellar hydrogen in the Galactic System with particular reference to the region within 3 kpc of the Centre.' Proc Nat Acad Sci 46 1 1960

Rowan-Robinson, M (1977) 'On the unity of activity in galaxies' Astrophys J 213 635 1977

Ryan, M.P. (1972) 'Is the existence of a galaxy evidence for a black hole at its centre ?' Astrophys J (Lett) 177 L79 1972

Salpeter, E.E. (1955) 'The luminosity function and stellar evolution' Astrophys J 121 161 1955

Salpeter, E.E. & Wagoner, R.V. (1971) 'Structure of relativistic discs' Astrophys J 164 557 1971

Sanders, R.H. (1970) 'The effects of collisions in dense stellar systems' Astrophys J 162 791 1970

Sanders, R.H. & Bania, T.M. (1976) 'The observational effects of explosions in the nuclei of spiral galaxies' Astrophys J 204 341 1976

Sanders, R.H. & Lowinger, T. (1972) 'The distribution of mass in the Galactic Nucleus' Astron J 77 292 1972

Sanders, R.H. & Prendergast, K.H. (1974) 'The possible relation of the 3 kpc arm to explosions in the Galactic Nucleus' Astrophys J 188 489 1974

Sanders, R.H. & Wrixon, G.T. (1973) 'Observations of neutral hydrogen near the Galactic Centre II. The Nuclear Disc' Astron & Astrophys 26 365 1973

Sanders, R.H. & Wrixon, G.T. (1974) 'Gas kinematics near the Galactic Centre from CO observations' Astron & Astrophys 33 9 1974

Sanders, R.H., Wrixon, G.T., & Mebold, U. (1977) 'Neutral hydrogen in the Galactic Center Region: A preliminary survey with the 100 m telescope' Astron & Astrophys 61 329 1977

Sandquist, A. (1974) 'Lunar occultations of the Galactic Centre Region in H I, OH and H_2CO lines II. The continuum and the + 40 km s^{-1} feature.' Astron & Astrophys 33 413 1974

Sargent, W.L.W. (1972) 'The spectra of 80 galaxies in Markarian's second list and the space density of Markarian galaxies' Astrophys J 173 7 1972

Sargent, W.L.W., Young, P.J., Boksenberg, A., Shortridge, K., Lynds, C.R., & Hartwick, F.D.A. (1978) 'Dynamical evidence for a central mass concentration in the galaxy M 87' Astrophys J 221 731 1978

Saslaw, W.C. (1974) 'Theory of Galactic Nuclei' I.A.U. Symp. 58 305 1974

Saslaw, W.C. & de Young, D.S. (1971) 'On equipartition in galactic nuclei and gravitating systems' Astrophys J 170 423 1971

Schwarz, J., McCray, R., & Stein, R.F. (1972) 'Formation of clouds in a cooling interstellar medium' Astrophys J 175 673 1972

Scoville, N.Z. (1972) 'Kinematics of molecular clouds near the Galactic Centre' Astrophys J (Lett) 175 L127 1972

Scoville, N.Z., & Solomon, P.M. (1974) 'Radiative transfer, excitation, and cooling of molecular emission lines (CO and CS)' Astrophys J (Lett) 187 L67 1974

Scoville, N.Z., Solomon, P.M., & Jefferts, K.B. (1974) 'Molecular clouds in the Galactic Nucleus' Astrophys J (Lett) 187 L63 1974

Shields, G.A., & Wheeler, J.C. (1978) 'Sustenance of a black hole in a galactic nucleus' Astrophys J 222 667 1978

Shklovsky, I.S. (1960) 'Radio galaxies' Sov Astron 4 885 1961

Shklovsky, I.S. (1962) 'On the nature of radio galaxies' Sov Astron 6 465 1963

Shklovsky, I.S. (1971) 'On the ejection of gas from active nuclei of galaxies' Astrophys Lett 10 5 1971

Shklovsky, I.S. (1975) 'Is the Galactic Nucleus a black hole ?' Sov Astron Lett 1 135 1975

Simónson III, S.C., & Mader, G.L. (1973) 'Motions near the Galactic Centre and the '3 - kpc arm' ' Astron & Astrophys 27 337 1973

Smith, F.G., O'Brien, P.A., & Baldwin, J.E. (1956) 'The radio source in the direction of the Galactic Centre' Mon Not Roy Astron Soc 116 282 1956

Spitzer, L. (1949) 'The temperature of interstellar matter III.' Astrophys J 109 337 1949

Spitzer, L. (1968) 'Dynamics of the Interstellar Gas' No. 28 of Interscience tracts on Physics and Astronomy

Spitzer, L. (1969) 'Equipartition and the formation of compact nuclei in spherical stellar systems' Astrophys J (Lett) 158 L139 1969

Spitzer, L. (1971) 'Dynamical evolution of dense star systems' Nuclei of Galaxies p. 443, 1971

Spitzer, L., & Hart, M.H. (1971) 'Random gravitational encounters and the evolution of spherical stellar systems II: Models' Astrophys J 166 483 1971

Spitzer, L. & Saslaw, W.C. (1966) 'On the evolution of galactic nuclei' Astrophys J 143 400 1966

Spitzer, L. & Schwartzschild, M. (1951) 'The possible influence of interstellar clouds on stellar velocities' Astrophys J 114 385 1951

Spitzer, L. & Thuan, T.X. (1972) 'Random gravitational encounters and the evolution of spherical systems IV. Isolated systems of identical stars' Astrophys J 175 31 1972

Stothers, R. (1966) 'Quasars as the origin of primordial matter in a steady-state universe' Mon Not Roy Astron Soc 132 217 1966

Sturrock, P.A. (1966) 'A model of quasi-stellar radio sources' Nature 211 697 1966

Sturrock, P.A. & Barnes, C. (1972) 'Activity in galaxies and quasars' Astrophys J 176 31 1972

Talbot, R.J., & Arnett, W.D. (1973) 'The evolution of galaxies III. Chemical evolution coefficients' Astrophys J 186 51 1973

Tammann, G.A. (1974) 'Statistics of supernovae' In Supernovae and Supernova Remnants (Ed: C.B. Cosmovici) p. 155, 1974

Tammann, G.A. (1977) 'A progress report on supernova statistics' In Supernovae (Ed: D. Schramm) p. 95, 1977

Tayler, R.J. (1976) 'A possible model for blue compact galaxies' Mon Not Roy Astron Soc 177 39 1976

Thompson, R.I., Lebofsky, M.J., & Rieke, G.H. (1978) 'The 2 - 2·5 micron spectrum of NGC 1068: Detection of extragalactic molecular hydrogen' Astrophys J (Lett) 222 L49 1978

Thorne, K.S. (1966) 'Relativistic stellar structure and dynamics' In: High Energy Astrophysics Volume III (Ed: De Witt, Schatzman & Véron) 1966

Thorne, K.S. (1974) 'Disk accretion onto a black hole II. Evolution of the hole' Astrophys J <u>191</u> 507 1974

Thorne, K.S., & Braginsky, V.B. (1976) 'Gravitational wave bursts from the nuclei of distant galaxies and quasars: Proposal for detection using Doppler tracking of interplanetary spacecraft' Astrophys J (Lett) <u>204</u> L1 1976

Tinsley, B.M. (1972) 'Galactic evolution: Program and initial results' Astron & Astrophys <u>20</u> 383 1976

Tinsley, B.M. (1976) 'Chemical evolution in the solar neighbourhood II. Statistical constraints, finite stellar lifetimes, and inhomogeneities' Astrophys J <u>208</u> 797 1976

Tinsley, B.M. (1977) 'Masses of supernova progenitors' Supernovae (Ed: D. Schramm) p.117, 1977

Tinsley, B.M. (1978) 'Evolutionary synthesis of the stellar population in elliptical galaxies II. Late M giants and composition effects' Astrophys J <u>222</u> 14 1978

Tremaine, S.D., Ostriker, J.P., & Spitzer, L (1975) 'The formation of the nuclei of galaxies I. M 31' Astrophys J <u>196</u> 407 1975

Ulrich, M-H (1978) 'Observations of the outflow of gas from the nucleus of NGC 253 and its implications for the stellar population of the nucleus' Astrophys J <u>219</u> 424 1978

Van den Bergh, S. (1975) 'Stellar populations in galaxies' Ann Rev Astron & Astrophys <u>13</u> 217 1975

Van den Bergh, S. (1976) 'The post-eruptive galaxy NGC 5128 = Centaurus A' Astrophys J <u>208</u> 673 1976

Van der Kruit, P.C. (1971) 'Attempt to explain the motions of the gas in the central region of the Galaxy by explosive events in its nucleus' Astron & Astrophys <u>13</u> 405 1971

Van der Kruit, P.C. (1973) 'High resolution radio continuum observations of bright spiral galaxies at 1415 M Hz : A general discussion' Astron & Astrophys 29 263 1973

Van der Kruit & Allen, R.J. (1976) 'The radio continuum morphology of spiral galaxies' Ann Rev Astron & Astrophys 14 417 1976

Van der Laan, H., & Perola, G.C. (1969) 'Aspects of radio galaxy evolution' Astron & Astrophys 3 468 1969

Von Hoerner, S., & Saslaw, W.C. (1976) 'The evolution of massive collapsing gas clouds' Astrophys J 206 917 1976

Wagoner, R.V. (1971) 'Production of Helium in massive objects' Highlights in Astronomy 2 301 1971

Wagoner, R.V. & Salpeter, E.E. (1972) 'Supermassive stars' External Galaxies and Quasi-stellar Objects (Ed: D.S. Evans) p.300 (I.A.U. Symp. 44 300 1972)

Walker, M.F. (1968) 'Studies of extragalactic nebulae V. Motions in the Seyfert galaxy NGC 1068' Astrophys J 151 71 1968

Weedman, D.W. (1976a) 'Luminosities of Seyfert galaxies and QSOs' Astrophys J 208 30 1976

Weedman, D.W. (1976b) 'Seyfert galaxies, quasars, and redshifts' Quat J Roy Astron Soc 17 227 1976

Werner, M.W., & Salpeter, E.E. (1969) 'Grain temperatures in interstellar dust clouds' Mon Not Roy Astron Soc 145 249 1969

Wolfe, A.M. (1974) 'On the steady flow of gas from the nuclei of Seyfert galaxies' Astrophys J 188 243 1974

Wolfe, A.M., & Burbidge, G.R. (1970) 'Black holes in elliptical galaxies' Astrophys J 161 419 1970

Wollman, E.R., Geballe, T.R., Lacy, J.H., & Townes, C.H. (1976) 'Spectral and spatial resolution of the 12·8 micron Ne II emission from the Galactic Centre' Astrophys J (Lett) 205 L5 1976

Wollman, E.R., Geballe, T.R., Lacy, J.H., & Townes, C.H. (1977) 'Ne II 12·8 micron emission from the Galactic Centre II' Astrophys J (Lett) 218 L103 1977

Wolstencroft, R.D., & Zealey, W.J. (1975) 'The peculiar galaxy NGC 1097' Mon Not Roy Astron Soc 173 51p 1975

Young, P.J. (1977) 'The black tide model of QSOs II. Destruction in an isothermal sphere' Astrophys J 215 36 1977

Young, P.J., Westphal, J.A., Kristian, J., Wilson, C.P., & Landaver, F.P (1978) 'Evidence for a supermassive object in the nucleus of the galaxy M 87 from SIT and CCD area photometry' Astrophys J 221 721 1978